METHODS OF COOPERATIVE ROUTING TO OPTIMIZE THE LIFETIME OF MULTI-HOP WIRELESS SENSOR NETWORKS

A Dissertation
Presented to
The Academic Faculty

by

Jin Woo Jung

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2013

METHODS OF COOPERATIVE ROUTING TO OPTIMIZE THE LIFETIME OF MULTI-HOP WIRELESS SENSOR NETWORKS

Approved by:

Dr. Mary Ann Ingram, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Ying Zhang
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Geoffrey Li
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Johan G. F. Belinfante
School of Mathematics
Georgia Institute of Technology

Dr. Raghupathy Sivakumar
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: January 9, 2013

To my beloved, Insun.

ACKNOWLEDGEMENTS

My time in Georgia Tech has been one of the most exciting periods in my life, and it has given me invaluable knowledge, countless pleasant memories, and precious experiences. I owe this to my family, friends, professors, colleagues, and Georgia Tech.

I wish to express my deepest gratitude to my PhD advisor, Prof. Mary Ann Ingram (Weitnauer). She has patiently encouraged my research by providing insightful suggestions and criticism, and also, she respected, believed and supported my ideas and approaches, which made all my PhD achievements possible. I will always remember her willingness to give her time so generously for my research and provide help for my personal matters. I am forever in debt to her, and it's been a great honor and privilege for me to work with her.

I also would like to thank prof. Raghupathy Sivakumar, Prof. Geoffrey Li, Prof. Ying Zhang, and Prof. Johan G. F. Belinfante for being my committee members. In particular, I am grateful to Prof. Raghupathy Sivakumar for generously accepting to be one of my references, Prof. Geoffrey Li for serving as my PhD proposal committee chair, Prof. Ying Zhang for attending my PhD proposal remotely, and Prof. Johan G. F. Belinfante for being my defense committee member even though he is retired.

My grateful thanks are also extended to Dr. Hee-Won Park (Vice President, Samsung Electronics) and the professors of Yonsei University, Prof. Hong-Yeop Song (my MS thesis advisor), Prof. Chungyong Lee (also a Georgia Tech PhD), and Prof. Young Yong Kim, for writing recommendation letters for my PhD study. I would also like to extend my thanks to Dr. Cheng Tan (Senior Manager, Apple) and John Avery (Group Manager, Panasonic) for giving me the internship opportunities.

I thank my awesome labmates at the Smart Antenna Research Lab (in the order of seniority), Dr. Guillermo Acosta, Alper Akanser, Lakshim, Dr. Aravind Kailas, Dr. Syed Ali Hassan, Yong Jun Chang, Steve William, Vivek Agate, Haejoon Jung, Jian Lin, and Van Nguyen, and visiting scholars, Mingxi Zhang, Nikolaj Marchenko, Wensi Wang, Jason

Hongzhi, and Gao Zhen, for spending time and making good memories with me; I also enjoyed discussing research topics with them. In particular, I would like to express my great appreciation to Dr. Aravind Kailas for his everlasting friendship and help and some of my Korean buddies, Seok-Chul Kwon, Yong Jun Chang, and Haejoon Jung, for their limitless support, encouragement and help.

I'd like to also thank all the members of the Yonsei University Alumni Association (YUAA) in Georgia Tech for their support, Dr. Won-Yeol Lee and Dr. Yeonsik Jeong for many helps and advices during the early period of my PhD study, Dr. Hyungseok Yu for giving me various helpful information, and the staff members, Pat Dixon and Cordai Farrar, for their help. Also, I deeply appreciate the financial support from the National Science Foundation (NSF) and Prof. Mary Ann Ingram.

My special thanks goes to my wife, Insun Ju, for her unlimited support during my PhD study. She has been there and is always there for me when I need her most, and she has been taking care of many things for me including house chores and cooking. Thanks to her sacrifice, I was able to concentrate on my PhD study, and I owe all my achievements to her. Also, I'd like to pay my gratitude to my family for the financial support during the early period of my PhD study and my dog, Angel (although he has not been so helpful in doing my PhD study), for always bringing laughter to me and my wife.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation presents methods of extending the network lifetime of multi-hop wireless sensor networks (WSNs) through routing that uses cooperative transmission (CT), referred to as cooperative routing. CT can have a signal-to-noise ratio (SNR) advantage over non-CT schemes through cooperative diversity and simple aggregation of transmit power, and one of its abilities is to extend the communication range of a wireless device using this SNR advantage. In this research, we use the range-extension ability of CT as a tool to mitigate the energy-hole problem of multi-hop WSNs and extend the network lifetime.

The main contributions of this research are (i) an analytical model for a cooperative routing protocol with a deployment method, (ii) cooperative routing protocols that can extend the network lifetime, and (iii) formulating the lifetime-optimization problem for cooperative routing. The analytical model developed in this research theoretically proves that, in a situation where non-CT routing cannot avoid the energy-hole problem, our CT method can solve the problem. Proactive CT (PROTECT), a CT method based on the analytical model, provides a very simple way of doing cooperative routing and can improve the lifetime of non-CT networks significantly. Residual-energy-activated CT (REACT), a cooperative routing protocol that uses the energy information of nodes, overcomes some of the limitations of PROTECT and can be applied to any existing non-CT routing protocol to improve the network lifetime. Using REACT and analytical approaches, we also show that cooperative routing can be beneficial in multi-hop energy-harvesting WSNs. By formulating and solving the lifetime-optimization problem of cooperative routing, which requires a much more sophisticated formulation than that of non-CT routing, we explore the optimal lifetime bounds and behaviors of cooperative routing. Finally, we study and design online cooperative routing methods that can perform close to the optimal cooperative routing.

CHAPTER 1

INTRODUCTION

In recent years, wireless sensor networks (WSNs) have been extensively studied by many researchers. WSNs consist of devices (nodes) that have sensing and communication capabilities, and the data collected from the nodes is usually gathered in one or more base stations (sink nodes), which are usually considered to have no energy constraint. The communication range of a wireless sensor node is limited, e.g., MICAz [1], a widely used off-the-shelf wireless sensor mote, has an indoor range of 20-30m and an outdoor range of 75-100m. If every sensor node can have a direct connection to its destination (sink node), that is, when the distance between any sensor node and its destination is within the communication range of the node, WSNs can be single-hop. Otherwise, WSNs have to be multi-hop, where some nodes have to relay other nodes' data. Multi-hop networks can be less expensive than single-hop networks because fewer sink nodes are required to cover a large area, and also, there can be situations where using a multi-hop network is inevitable because of the limited communication range of a node. Therefore, we consider multi-hop environments in this dissertation, and, among the many issues of multi-hop WSNs, extending the lifetime of WSNs through network-layer approaches is studied in this research.

Sensor nodes usually run on batteries, and, because (i) the battery power is limited and (ii) changing the batteries of sensor nodes is not always possible, extending the network lifetime of WSNs has been a critical issue [3], [4], [5]. Instead of using conventional routing approaches to extend the lifetime of WSNs, this research explores routing methods that use cooperative transmission (CT); CT is a method of improving the received signal-to-noise ratio (SNR) of a single-antenna communication device by virtually forming a multiple-antenna system using multiple transmitting devices. Optimizing the network lifetime of multi-hop WSNs using routing methods that utilize CT, referred to as "cooperative routing," is the main topic of this dissertation.

We optimize the network lifetime by overcoming the "energy-hole" problem [19] of multi-hop WSNs through cooperative routing. The energy-hole problem can be described as the situation when the batteries of nodes near the sink deplete early because these nodes are heavily burdened with traffic from the rest of the network. Since CT requires multiple transmitters to transmit a single packet, whereas non-CT (that does not use CT) requires only one transmitter, CT may consume more energy than non-CT, and therefore, in order to solve the energy-hole problem and extend the network lifetime through CT, one must devise a way to use CT wisely. This research finds the ways to optimally use CT by developing an analytical model and formulating the lifetime-optimization problem for CT. Also, to cope with the situations where the optimal solution cannot be used, online cooperative routing methods are designed, analyzed, and evaluated.

1.1 Research Contributions

Our research contributions are as follows:

- **Analytical Model for CT**

  One contribution of this research is the development of an analytical model for avoiding the energy hole using CT [75], [76]. Additional contributions made by developing the analytical model are (i) proposing a simple CT method, referred to as PROTECT, that can extend the network lifetime, (ii) showing that our CT method can solve the energy-hole problem under the uniform distribution of nodes; this problem is known to be unavoidable for the non-CT case in [21] and [25], and (iii) providing the mathematical expression for the expected extended lifetime when our CT method is used, which is very useful because the lifetime performance can be obtained without running network simulations.

- **REACT**

  To cope with the limitation of PROTECT (that is, requiring a specific node deployment), we have designed a new cooperative routing protocol, referred to as REACT,

which uses the energy information of nodes [77]. REACT can be applied to any existing non-CT routing protocol and significantly improve the network lifetime. Also, REACT requires only the information of the one-hop neighbors to utilize a CT link, which makes the routing protocol simple and feasible.

- **Cooperative Routing in Energy-Harvesting WSNs**

  We show the advantages of using cooperative routing in energy-harvesting WSNs (EH-WSNs) through (i) a simple analysis and (ii) developing a method of determining the supportable service of EH-WSNs that use the optimal non-CT routing [78]. The benefits of using cooperative routing in EH-WSNs are shown by comparing the performance of the REACT protocol with that of the optimal non-CT routing case.

- **Lifetime-Optimization Problem of Cooperative Routing**

  Motivated by the superior performances of REACT, the lifetime-optimization problem of cooperative routing is formulated [79], [80], which is another key contribution of this research. By using the formulated problem, the optimal lifetime and behavior of cooperative routing are analyzed, and we also identify important design parameters of the optimal cooperative routing protocol, which can help simplify the overall protocol design [79].

- **Online Cooperative Routing Methods**

  The last contribution of this research is designing online cooperative routing methods that can perform close to the optimal cooperative routing. Unlike the existing online routing methods that provide little or no analytical justification, we fully justify the proposed online methods by defining and utilizing the design criteria for online routing.

1.2 Dissertation Outline

The remainder of this dissertation is organized as follows. In Chapter 2, we review some of the related research topics. The common terms, definitions and assumptions that are used throughout the entire dissertation are also summarized in Chapter 2. In Chapter 3, an analytical model for CT is developed, and we introduce PROTECT. Chapter 4 presents REACT, and the performances of the REACT protocol under various situations are verified and evaluated using network simulations. The advantages of using cooperative routing in EH-WSNs are explored in Chapter 5, where we show that cooperative routing can provide better services compared to non-CT routing in EH-WSNs through analytical approaches and simulations. In Chapter 6, the lifetime-optimization problem of cooperative routing for multi-hop WSNs is formulated, and we evaluate the formulated problem for various cases. By analyzing the evaluation results, we also determine the key routing behaviors and design guidelines of cooperative routing in Chapter 6. The online cooperative routing methods that can perform close to the optimal cooperative routing are studied and designed in Chapter 7. Finally, in Chapter 8, concluding remarks are provided along with the directions for future research.

CHAPTER 2

BACKGROUND

In this chapter, we review some of the related research topics, and we make common definitions and assumptions. We first summarize non-CT lifetime-extension routing approaches in Section 2.1. Then, in Section 2.2, the energy-hole problem of multi-hop WSNs is discussed. We overview CT and cooperative routing in Section 2.3. Finally, the common terms, definitions and assumptions that are used throughout the entire dissertation are introduced in Section 2.4.

2.1 Non-CT Routing Approaches to Extend the Network Lifetime of Wireless Networks

In multi-hop wireless networks, there can be many routes from a source to a destination, and since a chosen route has an impact on the energy consumption of each node involved in the route, many authors have proposed routing methods that can minimize the total energy consumption of the route [26], [27], [28]. This type of routing, which will be referred to as the minimum total energy (MTE) routing in this dissertation (following [34]), may guarantee the minimum energy consumption of the entire network, however, as is pointed out in [29], this is not a desirable way to extend the network lifetime because the nodes in the MTE route are used more frequently than the others resulting in early death for the frequently used nodes.

Energy-aware (or power-aware) routing is an advanced version of MTE routing. Energy-aware routing not only considers the energy consumption of the route but also the remaining or residual energy of each node to extend the network lifetime [29], [30], [31], [32], [33], [34], [35]. Maximizing the network lifetime through off-line (meaning that the future generation of traffic can be known in advance) routing is addressed in [31]. When the future generation of traffic is unknown, which is more general case than the off-line case,

an "online" routing method is required, and various online energy-aware routing protocols are proposed in [29], [32], [33], [34], and [35]. Note that, because online routing has no information on how the traffic will be generated, it is designed to behave reasonably; the route is selected in a way that the total energy consumption is low while the nodes with low residual energy are avoided. Also, because of the uncertainty of the future generation of traffic, there is no guarantee that the performance of an online routing protocol will always be close to the optimal performance (here, the optimal performance is the performance of the off-line case).

Two prominent online energy-aware routing works are the flow augmentation (FA) algorithm in [34] and the capacity maximization (CMAX) algorithm in [35]. The work in [34] (originally published in [31]) provides simulation results that compare the lifetime performance of the FA algorithm with that of the MTE routing, conditional max-min battery capacity routing (CMMBCR) in [33], max-min $zP_{\min}$ algorithm in [32], and the optimal off-line case (based on the lifetime-optimization problem introduced in [31]) to justify the superiority of the FA algorithm. In [35], the authors present the competitive-ratio analysis to show the effectiveness of their CMAX algorithm. Because of the importance of the analysis and routing works in [31], [34], and [35], we briefly introduce them in the following subsections.

2.1.1 FA Algorithm [31], [34]

Let (i,j) denote the link between Node i and Node j. The FA algorithm uses the following link-cost metric for each link (i,j):

$$e_{ij}^{x_1} \cdot r_i^{-x_2} \cdot (E_i^{\mathrm{I}})^{x_3}, \tag{1}$$

where e_{ij} is the energy consumption associated with the link (i, j), r_i is the residual energy of Node i, E_i^{I} is the initial energy of Node i, and x_1, x_2, and x_3 are nonnegative weighting factors [31]. Then, the route from a source to a destination is chosen by selecting the set of links that minimizes the sum of link costs. Note that the method of selecting the

minimum-cost route is a widely used approach for online energy-aware routing [29], [30], [33], [35], and this method can be implemented using any existing shortest-path algorithm including the distributed Bellman-Ford algorithm [10]. As for the weighting factors, the authors recommended using x_2=x_3 based on their simulation results.

2.1.2 Lifetime-Optimization Problem of Non-CT Routing [31]

In [31], the authors formulated the lifetime-optimization problem for non-CT routing using linear programming (LP) [11]. Before introducing the formulation, we make several definitions.

We define the lifetime, T, of the network to be the time that the first node dies. We let A be the set of all nodes in the network, S_i be the set of neighbors of Node i, D be the set of destination nodes, and E be the set of nodes that are not energy constrained. We define e_{ij}^{TX} as the required energy for Node i to transmit a data unit from Node i to Node j, e_{ji}^{RX} as the required energy for Node i to receive a data unit coming from Node j, and Q_i as the information generation rate (data/time) of Node i ($Q_i > 0$). Note that since e_{ji}^{RX} is just circuit energy consumption, we use e^{RX} for the receiving energy instead of e_{ji}^{RX}. Finally, we define n_{ij} as the total number of data units transmitted from Node i to Node j until the lifetime T of the network.

Using the above definitions, the lifetime-optimization problem for non-CT routing can be stated as follows:

$$\text{Maximize } T$$

$$\text{s.t. } n_{ij} \geq 0, \qquad \forall i \in A, \forall j \in S_i, \tag{2}$$

$$\sum_{j \in S_i} e_{ij}^{\mathrm{TX}} \cdot n_{ij} + \sum_{j : i \in S_j} e^{\mathrm{RX}} \cdot n_{ji} \leq E_i^{\mathrm{I}}, \qquad \forall i \in A - E, \tag{3}$$

$$\sum_{j : i \in S_j} n_{ji} + T \cdot Q_i = \sum_{j \in S_i} n_{ij}, \qquad \forall i \in A - D, \tag{4}$$

where E_i^{I} is the initial energy of Node i defined in Section 2.1.1. (3) is the energy-constraint condition, which indicates that the total energy consumption of Node i during the network

lifetime T cannot exceed E_i^{I}. (4) is the data-conservation condition[1] of Node i, which indicates that the total incoming data should be equal to the total outgoing data. Note that $T \cdot Q_i$ is the number of data units generated by Node i during the lifetime of the network, and it is considered as the incoming data of Node i.

2.1.3 CMAX Algorithm [35]

Here, we use some of the variables defined in Section 2.1.1. The CMAX algorithm uses the following link-cost metric for each link (i,j):

$$e_{ij} \cdot (\lambda^{\alpha_i} - 1), \tag{5}$$

where α_i is $(1-r_i/E_i^{\mathrm{I}})$, which is the fraction of the used energy of Node i, and λ is a constant. There is no fixed way to select λ, however, the authors showed that in a practical case where a message is never dropped if a node has sufficient energy to handle the message, the network capacity is relatively insensitive to λ as long as λ is large enough, and λ=100 is used in most cases [35]. Using (5), the minimum-cost route is determined and used.

The competitive-ratio analysis is also provided in [35] to show the effectiveness of the CMAX algorithm. The competitive ratio of an online algorithm is the worst case ratio of the capacity of the off-line algorithm to the capacity of the online algorithm, where the capacity is measured by the total length of messages successfully routed to the destination [35]. The authors in [35] conclude that the following competitive bound holds:

$$\frac{L(k)}{L_{opt}(k)} \geq \frac{1}{1 + 2\log_2 \lambda}, \tag{6}$$

where $L(k)$ is the total length of messages successfully routed by CMAX till the arrival of message k, and $L_{opt}(k)$ is the total length of messages successfully routed by the optimal algorithm till the arrival of message k.

[1] In [31], (4) is called "flow-conservation" condition. We use the term "data conservation" in this dissertation because we are directly looking at the amount of data instead of flows as in [50].

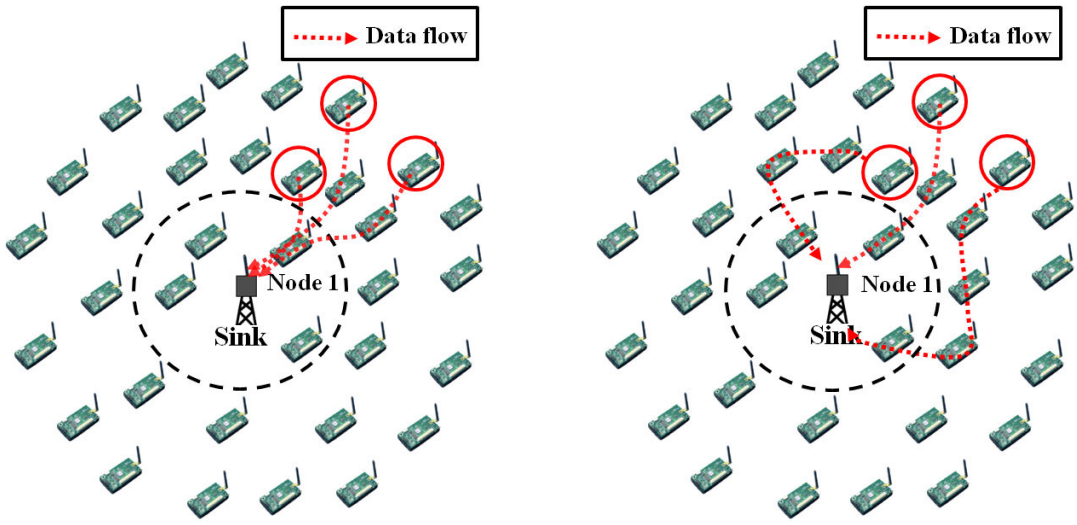
2.2 The Energy-Hole Problem

Multi-hop WSNs suffer from the so-called "energy-hole" problem [19], which can be described as the situation when the nodes around the sink nodes consume relatively more energy than the other nodes (because they have to relay the other nodes' data) and die early. To approach the problem analytically, most of the works that address the energy-hole problem consider the case where there is a single sink node in the network, referred to as "many-to-one network[2]" [19], [20], [21]. Note that, this many-to-one network is the worst case situation from the energy consumption point of view because, if there are multiple sink nodes, the traffic can be distributed to each sink node to better balance the energy consumption. When the energy hole is created in a many-to-one network, not a single packet can reach the sink node, and the whole network becomes useless, even though a majority of nodes still have enough energy to sense and send data. Because of this reason, the network lifetime is bounded by the lifetime of the nodes close to the sink node. The energy-hole problem is illustrated in Figure 1a, where it is shown that Node 1, which is one hop away from the sink node, is highly burdened (used three times).

The energy-aware routing protocols introduced in Section 2.1 are ineffective in solving the energy-hole problem because the fact remains that the nodes close to the sink node have to relay all the data. This is illustrated in Figure 1b; the energy-aware routing may avoid using Node 1, however, for each data flow, one of the nodes one hop away from the sink node must be used to deliver the data to the sink node.

The existing solutions for the energy-hole problem that involve the network-layer protocol are (i) the non-uniform node deployment [25], [21], [24] and (ii) the mobile-node strategy [22], [23], as illustrated in Figure 2. In the non-uniform node deployment, additional nodes are placed in the area close to the sink node, and routes are selected so that

[2]The case where a network has multiple sink nodes can be divided into multiple many-to-one networks when a source node is allowed to send its data to one of the sink nodes only.



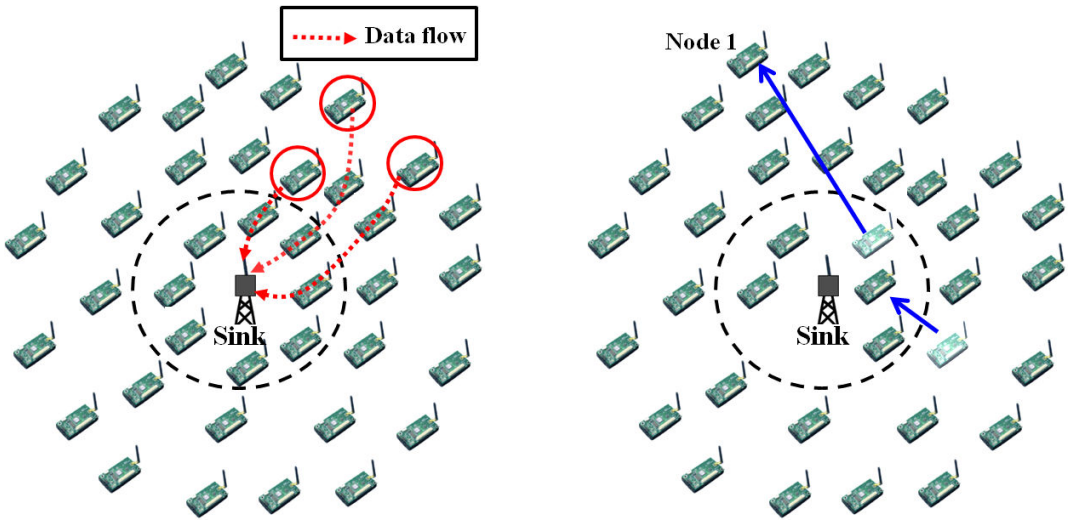
(a) The energy-hole problem when the minimum-hop-based routing is used.

(b) The energy-hole problem when the energy-aware routing is used.

Figure 1: Illustrations of the energy-hole problem in many-to-one network. The sink node is located at the center of the network, and the nodes in the dashed circle are the nodes that are one hop away from the sink nodes. Three data flows (routes) from three different source nodes are indicated using dotted arrows.

nodes can evenly consume the energy. The downside of this strategy is that it can drastically increase the cost of deployment because of the required additional nodes [21], or, from a different perspective, when the number of available nodes in the network is fixed, the strategy decreases the sensing/coverage area of the network because the additional nodes placed near the sink node could be used to cover other areas. The mobile-node strategy determines the movement of mobile nodes and routes (sending packets via mobile node) to mitigate the energy-hole problem [22], [23]. However, as noted in [22], mobile nodes may be hard to operate in certain environments such as under a bridge, on water, and in an unpaved area.

Before moving on, we provide one important result of [21]. In [21], a homogeneous wireless sensor network with all sensors having the same maximum transmission range and data-generation rate is assumed. In this type of WSNs, the authors showed that to at least partially balance the energy consumption, the non-uniform deployment is required. Moreover, even with the non-uniform deployment, the authors proved that completely balancing



(a) The non-uniform node deployment. More nodes are placed close to the sink node to distribute the traffic.

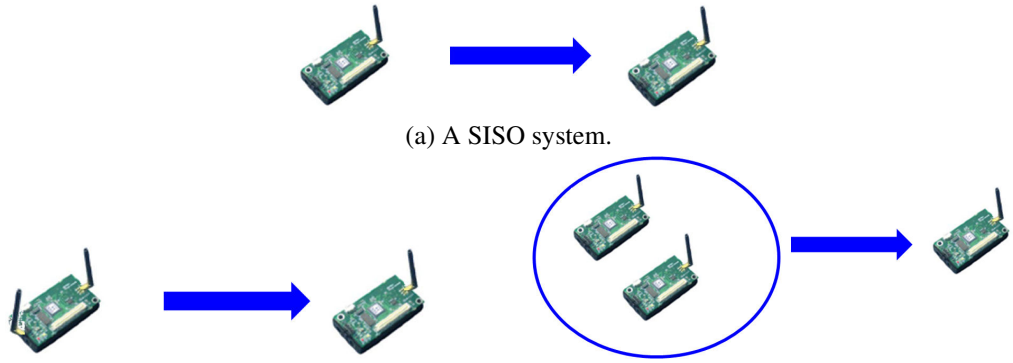
(b) The mobile-node strategy. Node 1 can move far away from the sink node, and another node can replace Node 1's role.

Figure 2: Illustrations of the solutions for the energy-hole problem of multi-hop WSNs.

the energy consumption of all nodes in the network is impossible because the nodes placed in the outmost area only have to send their data, whereas the other nodes have to (i) receive and send the data from the outmost area and (ii) send their own [21].

2.3 Cooperative Transmission and Cooperative Routing

CT [12] is a mixture of a communication protocol and a physical-layer combining scheme that can improve the communication quality of single-antenna communication devices. A transmitting node that uses CT shares its data packet with neighboring single-antenna nodes, and then, the collection of these nodes can transmit the packet to the intended receiver, thereby creating a virtual multiple-input-single-output (VMISO) system. The intended receiving node can use a physical-layer combining scheme to get diversity and array gains, which give CT an SNR advantage over the traditional single-input-single-output (SISO) case (non-CT case). This SNR advantage can be used to save transmit power, increase data rate, and extend the communication range. Figure 3 illustrates SISO, MISO, and VMISO systems.



(a) A SISO system.

(b) A 2x1 MISO system where the transmitting device has two antennas.

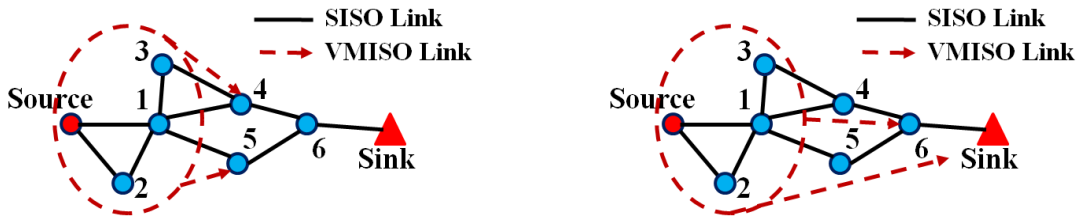
(c) A VMISO system where two single-antenna devices form 2x1 MISO system virtually.

Figure 3: Illustrations of SISO, MISO, and VMISO systems.

A VMISO system is particularly useful when a wireless device cannot have multiple antennas because of the hardware limitation (size, cost, etc.). If multiple devices simultaneously send the same data using the same channel, the multiple data transmissions may interfere with each other, which is undesirable. To avoid this problem, devices that participate in a VMISO communication may transmit in orthogonal channels, and the orthogonality can be achieved using time [12], [13], [14], frequency [52], and space-time coding [15], [16], [17].

Cooperative routing uses CT to improve the performance of routing. Here, we introduce two important CT approaches that are related to our research: (i) power-saving CT and (ii) range-extension CT. The main difference between the two on how a VMISO link is formed is illustrated in Figure 4. In the case of power-saving CT, a VMISO link is established between two nodes in a SISO link (more generally, at least one of the VMISO transmitters has a SISO connection with the VMISO receiver) [39], [40], [41], [42], [43], [44], [45] [46]. If we consider the case where both VMISO and SISO links have the same destination, because of the SNR advantage of a VMISO system over a SISO system, the required transmit power for each VMISO transmitter is lower than that of the single SISO transmitter, and, if the overall energy consumption of using the VMISO link is lower than that of using the

SISO link, power-saving CT can be used for better energy efficiency. Range-extension CT extends the one-hop SISO range of a node through the SNR advantage, and a new extended link can be established [47], [48], [49] (the link is "new" because there is no direct SISO connection between any of the VMISO transmitters and the VMISO receiver). Through the extended communication range, range-extension CT can increase routing choices and reduce the number of hops to the destination [47], [49], [48]. Note that, unlike power-saving CT, range-extension CT intentionally consumes more energy than non-CT to extend the SISO communication range of a node. [47] proposed a cross-layer framework for range-extension CT, which enables a node to gather cooperating nodes and utilize the extended VMISO link.



(a) Power-saving CT. (b) Range-extension CT.

Figure 4: Power-saving CT vs. range-extension CT. Four nodes (source node and Nodes 1-3) in the dashed circle are VMISO transmitters.

Many cooperative routing protocols that utilize power-saving CT focus on finding the energy-efficient (or power-efficient) route [40], [42], [37], [38], [45], [46], which is not different from the MTE routing approach discussed in Section 2.1. As MTE routing is not a desirable way to extend the network lifetime, so is finding the energy-efficient route by using power-saving CT because, as we have mentioned in Section 2.1, some nodes are used more frequently than the others resulting in early death for frequently used nodes. The works in [41], [43], and [44] consider the remaining lifetime of each device when forming a cooperative route using power-saving CT. Therefore, these works are suitable for extending the network lifetime as long as the energy efficiency can be achieved through power-saving CT. However, as discussed in [18], using power-saving CT may not be energy

efficient when the distance between two communicating nodes is not large enough or when the circuit energy is a large component of the total energy budget, as is often the case with sensor nodes [1]. When the overall energy efficiency cannot be achieved through VMISO links, the cooperative routing that relies on power-saving CT will always choose SISO routes, and no benefit can be obtained. Therefore, the cooperative routing that relies on power-saving CT cannot guarantee the lifetime extension of multi-hop WSNs.

The network-lifetime optimization using cooperative routing has been addressed by several authors [39], [41], [43], [49]. The works in [39] and [49] (and also many CT-based routing works [38], [40], [37], [45], [42]) ignore the circuit energy consumption, which oversimplifies the problem or makes their approaches incorrect. The works in [41] and [43], even though they consider the circuit energy consumption, may not be suitable for the lifetime optimization of multi-hop WSNs because they rely on power-saving CT, which, as mentioned above, cannot guarantee the lifetime extension of multi-hop WSNs.

The diversity-combining techniques that are essential in obtaining SNR advantages for CT are not new [8], [6], [7]. Also, the physical-layer implementation for CT has been done in practice. In [53] and [14], CT was successfully implemented and demonstrated, and, in [54], the authors have implemented various diversity-combining schemes for cooperative communication using off-the-shelf wireless sensor motes.

2.4 Common Terms, Definitions and Assumptions

In this section, we introduce common terms, definitions and assumptions that are used throughout the entire dissertation.

2.4.1 Terms and Definitions

Neighbors of a node are the ones that are within SISO communication range of the node. "Cooperative routing" is the routing that uses CT, and we use the term "non-CT" for the methods that rely on SISO communications only. A "cooperative route" is the route established using cooperative routing. The network that does not use CT is referred to as the

"non-CT network," and the network where both CT and non-CT are supported is referred to as the "CT network." In a CT network, a node, when it has a data to be transmitted, can either "do CT" by doing a VMISO communication with its selected neighbors or "do non-CT" by sending its data to one of its neighbors (a SISO communication). If a node decides to do CT, it becomes a "CT initiator" (or just "initiator"). *Cooperators* of a node are the neighbors of the node that are selected by the node to do CT. An initiator should share its data with its cooperators, and this is called the "CT sharing." When the initiator also participates in the VMISO communication, which is always the case in this dissertation, we use the term "cooperating nodes" to indicate the nodes that do CT, which include the initiator and the cooperators of the initiator. If a non-CT route needs to be established and used before a cooperative route is formed, we call the non-CT route the "primary route" [47][3]. Also, the non-CT routing scheme that is used to establish the primary route of cooperative routing is called "primary routing" scheme.

The total number of cooperating nodes is denoted by N_c. The diversity gain is a monotonically increasing function of N_c, and we denote this gain by $G(N_c)$. The maximum number of cooperating nodes is denoted by $N_c^{\max}$, and the maximum number of orthogonal diversity channels is denoted by N_d.

2.4.2 Assumptions

In this dissertation, we look at the problem from the network-layer perspective, and we use the network-layer approaches (i.e., routing protocols) to solve the problem. Using the techniques of other layers such as data aggregation and compression are not considered for both non-CT and CT networks.

We consider multi-hop WSNs, where the maximum transmission range of a wireless sensor device is not long enough to cover the entire network. Sensed data is gathered at sink nodes, and when there are multiple sink nodes, a source node needs only to send its

[3] [39], [42], and [41] are some of the CT works that rely on a non-CT primary route.

data to one of the sink nodes. A sink node has unlimited energy and resources[4] so that it can communicate with any of the nodes directly even without CT (following [20]). Nodes except for sink nodes are energy constrained.

In CT networks, if there are no prearranged cooperators (which is usually the case), the initiator is in charge of selecting the cooperators. The initiator is also in charge of sharing its data with the selected cooperators and doing the VMISO communication. The initiator always participates in the VMISO communication. The initiator and its cooperators transmit data to the VMISO receiver in orthogonal channels. The maximum number of cooperating nodes ($N_c^{\max}$) cannot exceed the maximum number of orthogonal diversity channels (N_d), and therefore, $N_c^{\max} \leq N_d$. Also, an initiator can select up to $N_c^{\max} - 1$ cooperators, and $2 \leq N_c \leq N_c^{\max} \leq N_d$ (note that N_c=1 is the non-CT case).

A node can successfully decode-and-forward a packet without an error when its received SNR is greater than or equal to a modulation-dependent threshold. Any node within the maximum transmission range of the transmitter will have high enough received SNR to decode the data without an error. For the physical layer, we assume a slowly varying (i.e., remains same for the entire VMISO transmission) Rayleigh fading channel with log-normal shadowing, such that the shadowing standard deviation is 5dB and the path loss exponent is 2.9 (values from [53], which are similar to [51]). The modulation scheme is assumed to be binary phase-shift keying (BPSK). To get the cooperative diversity gain, $G(N_c)$, in this environment, we obtain array gain and cooperative diversity gain using Monte Carlo simulation for $N_c = 1$ to 4 with target bit error rate (BER) of 10^{-3}, and then, we get the pure cooperative diversity gain by subtracting the array gain part ($10 \log_{10} N_c$) according to the number of cooperators. The resulting values of the cooperative diversity gain for N_c=2, 3, and 4 are given in Table 1.

[4]Note that the sink node in WSNs is usually considered to be supplied by a constant power source (such as AC power supply) and have greater computational resources than any other device(s) [2]. If the sink node has limited energy, it may die earlier than any other nodes because it has to receive and process every packet, and we don't consider such case.

Table 1: Diversity gain (BPSK. BER = 10^{-3}).

N_c	2	3	4
$G(N_c)$ (dB)	7.5	9.23	9.98

CHAPTER 3

AN ANALYTICAL MODEL FOR COOPERATIVE ROUTING WITH A FIXED TRANSMIT POWER

3.1 Overview

In this chapter, by using CT, we intend to find the solution of the energy-hole problem that can be applied to the network having limited capabilities. More specifically, we assume that wireless sensor nodes have limited information (neither residual energy nor distance between nodes is available) and use the same transmit power (no power control). To achieve the goal, we develop an analytical model for avoiding the energy hole using CT. Using this analytical model, a CT method that can balance the energy consumption and extend the network lifetime is developed. Unlike most of the existing literature on the energy-efficient cooperative routing protocols [41], [43], [44] that rely on the simulation results to figure out how much lifetime can be extended, we derive the amount of the expected extended lifetime when our CT method is used through this analytical model. Using our proposed CT method, the energy consumption can be perfectly balanced even with the uniform distribution of nodes, and therefore, our method can solve the energy hole problem under the uniform distribution, which is shown to be unavoidable for the non-CT case [25], [21]. We note that our solution introduced in this chapter is different from the two existing solutions introduced in Section 2.2 because we neither place more nodes per area nor require mobile units; instead, our method uses CT.

The remainder of this chapter is organized as follows. Section 3.2 introduces the models and assumptions that are used in this chapter. In Section 3.3, we develop the analytical model for CT and cooperative routing method. The network simulation results for our CT method are presented in Section 3.4. Some of the practical concerns are discussed in Section 3.5, and we summarize this chapter in Section 3.6.

3.2 Assumptions and Network Model

In addition to the common assumptions introduced in Section 2.4.2, we make the following assumptions. The network is a many-to-one multi-hop wireless sensor network that has a single sink node. Every node has the same initial energy. The nodes all use the same transmit power for transmission, which gives the same maximum transmission range, denoted by $d_{\text{tx}}^{\max}$. Also, we ignore the energy cost of control packets for both non-CT and CT, which is usually considered to be smaller than that of the data packets and ignored for analytical purposes [19], [20], [21].

When forming a cooperative route, a non-CT primary route is formed first. Non-CT routing is based on the shortest (minimum) hop, and the lifetime of the network when the shortest-hop routing is used will be compared with the case when our CT method is used. As mentioned in Section 2.2, the energy-aware routing protocols cannot successfully mitigate the energy-hole problem for many-to-one network, and this lets us focus on the shortest-hop routing, which is used in other papers that analyze the energy-hole problem [21], [20].

Our CT method utilizes the extended range obtained by CT and tries to form a VMISO link between cooperating nodes and the sink node. The mathematical expression for the extended range of CT, denoted by d_{ext}, is explained in Appendix A. Using the values of the diversity gain in Table 1, we can express d_{ext} in terms of $d_{\text{tx}}^{\max}$ for N_c=2, 3, and 4, which is summarized in Table 2. Note that when the sink node has unlimited energy and resources, the link from the sink node to the cooperating nodes can be formed using a SISO (non-CT) communication. Therefore, the diversity-combining technique to gain cooperative diversity needs to be applied to the sink node only, and the nodes other than the sink node do not require complicated hardware. We assume that the orthogonal diversity channels of CT are obtained by time [12].

We define C_i as the area of a circle with its center at the location of the sink node and a radius of $i \times d_{\text{tx}}^{\max}$, where $i \geq 1$ and $C_0 = 0$ as shown in Figure 5. The network outer boundary

Table 2: The extended range through CT (d_{ext}).

N_c	2	3	4
d_{ext}	$2.3d_{\text{tx}}^{\max}$	$3.04d_{\text{tx}}^{\max}$	$3.56d_{\text{tx}}^{\max}$

is C_L, a circle with radius $L \cdot d_{\text{tx}}^{\max}$, with its center at the sink node. We let $A_i = C_i - C_{i-1}$ be the area between two circle boundaries. We also refer to this A_i as the "i-th level" or "Level i." As in [20], we assume that the nodes in adjacent A_i's can communicate with each other. We also assume that the nodes in A_{i-1} equally share the traffic coming from A_i, which is the case when nodes are deployed uniformly or the node density is high [20]. Any node in the network except for the sink node can be a source at random.

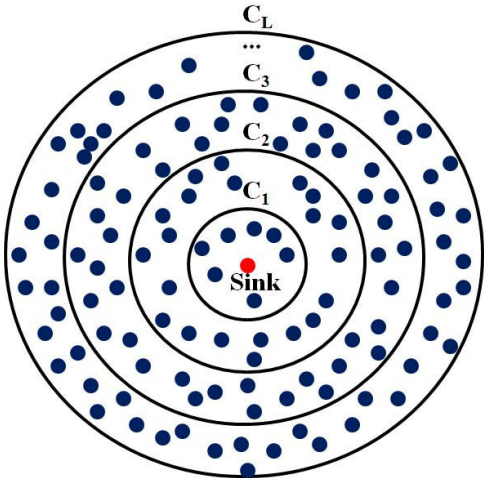


Figure 5: A circular-shaped network with size C_L.

3.3 The Analytical Model and Cooperative Routing Method

In this section, we develop the analytical model and CT method for avoiding the energy-hole problem. We briefly discuss our CT approach and define variables required for our analytical model in Section 3.3.1. In Section 3.3.2, we derive the analytical model for CT assuming a circular-shaped network (Figure 5) and the uniform distribution of nodes, by which we mean that the number of nodes per unit area is equal, across the network. Since

[21] and [25] showed that balancing the energy consumption is impossible for the circular-shaped network with the uniform distribution, by using the same type of the network, we show that the energy balancing can be achieved when our CT method is used. In Section 3.3.3, we present our deployment method and how to adapt our algorithm to various shapes of networks other than circular followed by our proposed cooperative routing protocol.

3.3.1 Our CT Approach and Variable Definitions

The main idea of our cooperative routing strategy is to balance the load of the nodes by using range-extension CT, which is illustrated in Figure 6. If the nodes in A_2 use CT and the extended range when CT is used is more than twice the single-hop distance of non-CT ($d_{\text{ext}} \geq 2d_{\text{tx}}^{\max}$), then the nodes in A_2 can reduce the burden of the nodes in A_1 by communicating directly to the sink node using CT. In general, if the nodes in A_i use CT and the range extended by CT is more than i times of the non-CT range ($d_{\text{ext}} \geq i \cdot d_{\text{tx}}^{\max}$), the nodes in A_j ($j < i$) can reduce their burdens. When CT is done by the nodes in M level(s) (starting from A_2 to A_{M+1}), it will be referred to as "M-Level" CT. M can be thought of as the maximum number of levels being "hopped over" by CT.

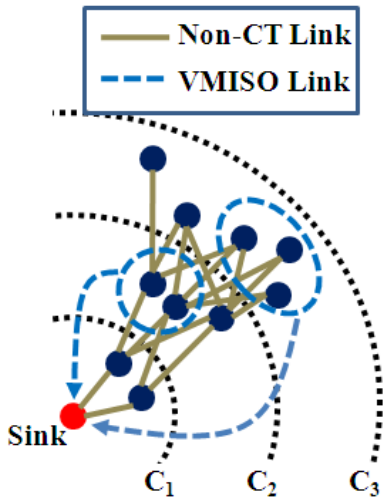


Figure 6: The illustration of using range-extension CT showing non-CT (SISO) links between adjacent levels and VMISO links between cooperating nodes and the sink node.

To define the network lifetime, the notion of "task" used in [20] is adopted. As in [20],

one task involves transmitting a packet of sensed data all the way to the sink node. The number of successfully performed tasks matches with the number of packets that successfully reach the sink node, and this value is widely used as the network lifetime [35], [36]. Therefore, the number of tasks that can be performed before the network is completely disconnected is considered as the network lifetime. Here, connectivity is determined by the number of nodes that can reach the sink node. We assume that tasks are well scheduled so that each task can be performed one after the other incurring no collision.

Following [20], we define the following variables. Let T_{nc} be the number of tasks performed during the network lifetime when non-CT is used and n_i be the number of the nodes in the area A_i. Since the nodes in A_1 need to take care of all the tasks in the network, the expected number of tasks per node in A_1 is T_{nc}/n_1. The nodes in A_i ($i \geq 2$) must relay all the tasks originating from A_k ($k > i$), but do not have to relay the ones from A_j ($j < i$). The total number of tasks originating from the nodes in A_i can be expressed as $\frac{n_i}{n} T_{nc}$ where n is the number of the nodes in the entire network ($n = \sum_{i=1}^{L} n_i$). Therefore, the total number of the tasks, T_i, that nodes in A_i should handle is

$$T_i = T_{nc} - \frac{T_{nc}}{n} \left(\sum_{k=1}^{i-1} n_k \right). \tag{7}$$

Then, the expected number of tasks per node in A_i is $\frac{T_i}{n_i}$. Note that $T_1 = T_{nc}$, which means that the nodes in A_1 should handle all the tasks till the lifetime of the network, and the network lifetime is bounded by the lifetime of the nodes in A_1.

We denote by T_c, the number of tasks performed during the lifetime of the network when CT is used. Also, we denote by $N_{c,i}$, the minimum number of the cooperating nodes in A_i required to directly reach the sink node ($i \geq 2$). Note that $N_{c,i}$ depends on CT's ability to extend the non-CT (SISO) range, and $N_{c,i}$ should be large enough to satisfy $d_{\text{ext}} \geq i \cdot d_{\text{tx}}^{\max}$. As we have explained earlier, the nodes in A_i ($i \geq 2$) can reduce the burden of the nodes in A_j ($j < i$) by using CT, and the number of tasks for which the nodes in A_i do CT will be denoted by $T_{c,i}$.

In the case of the uniform distribution, we define the node density as ρ, which leads to $\rho = \frac{n_i}{A_i}$ when nodes are uniformly deployed. We define E_i as the total energy consumed by a node in A_i when the network is dead. Also, E_{TX} and E_{RX} denote the energy consumptions for transmitting and receiving a packet (task), respectively, and E_{relay} denotes the energy consumption for relaying ($E_{\mathrm{relay}} = E_{\mathrm{TX}} + E_{\mathrm{RX}}$). Since we assume a fixed transmit power, E_{TX} is fixed, and we define a fixed positive variable η as $\eta = E_{\mathrm{RX}}/E_{\mathrm{TX}}$. Then, E_{relay}=$(1 + \eta)E_{\mathrm{TX}}$.

3.3.2 The Analytical Model for M-Level CT

Because of its relative simplicity, we first investigate the 1-Level CT ($M = 1$), in which CT is done only by the nodes in A_2. From the definition in Section 3.3.1, CT can perform T_c tasks, and the nodes in A_2 will do CT for $T_{c,2}$ tasks. In the case of non-CT, the nodes in A_1 should be involved in all tasks, however, in the case of CT, $T_{c,2}$ tasks are handled by the nodes in A_2, and the nodes in A_1 take care of only $T_c - T_{c,2}$ tasks. Therefore, a node in A_1 is involved in handling $(T_c - T_{c,2})/n_1$ tasks. Note that when a node is a source, it consumes only transmitting energy, E_{TX}, whereas, when a node is a relay, it consumes E_{relay}. Since each node in the network generates T_c/n tasks on average till the lifetime ends, this is the average amount of "source" tasks generated by a node in A_1. Therefore, the expected number of tasks that a node in A_1 has to relay is $(T_c - T_{c,2})/n_1$ - T_c/n, and the amount of energy that a node in A_1 consumes (E_1) is

$$E_1 = \left(\frac{T_c - T_{c,2}}{n_1} - \frac{T_c}{n}\right) \cdot E_{\mathrm{relay}} + \frac{T_c}{n} E_{\mathrm{TX}}. \tag{8}$$

Next, we obtain E_2, the total energy consumed by a node in A_2. Regardless of using CT or not, the actual number of tasks that the nodes in A_2 need to take care of remains the same (which is T_2 in (7)). However, when CT is used, the energy consumption of the nodes in A_2 per task will increase because cooperation involves the usage of more than one node. To account for the additional energy required for cooperation, we define the number of *virtual tasks* to be the total number of extra transmissions caused by CT. Virtual tasks do not add to

the total number of actual tasks (number of packets reached the sink node), but are required to take the additional energy consumption of CT into account. When handling the tasks in Level 2, one of the nodes in Level 2 is always used regardless of doing CT or not. Therefore, when $N_{c,2}$ nodes are cooperating, one of these nodes cannot be considered as facing the additional tasks because that node is used even in the case of non-CT. Because of this reason, the additional tasks caused by CT (virtual tasks) can be expressed as $(N_{c,2}-1)\times T_{c,2}$. Therefore, the nodes in A_2 should deal with $T_2 + (N_{c,2} - 1) \times T_{c,2}$ tasks, and the expected number of tasks per node in A_2 when CT is used is $(T_2 + (N_{c,2} - 1) \times T_{c,2})/n_2$. The average amount of source tasks generated by a node in A_2 is T_c/n, and therefore, we get E_2 as follows:

$$E_2 = \left(\frac{T_c - \frac{n_1}{n} T_c + (N_{c,2} - 1) T_{c,2}}{n_2} - \frac{T_c}{n} \right) \cdot E_{\text{relay}} + \frac{T_c}{n} E_{\text{TX}}. \tag{9}$$

When $E_2 > E_1$, the nodes in A_2 die first, and $E_2 < E_1$ creates the energy hole in A_1. Therefore, it is clear that the condition $E_2 = E_1$ balances the energy consumption so that the network lifetime can be optimally increased (the increased lifetime through our method is discussed later). Before solving $E_2 = E_1$, it is helpful to reduce the number of variables. Note that $A_i = \pi(i \cdot d_{\text{tx}}^{\max})^2 - \pi(i-1)^2 \cdot (d_{\text{tx}}^{\max})^2 = \pi(d_{\text{tx}}^{\max})^2(2i-1)$. The uniformity of node distribution gives $\rho = n_1/\{\pi(d_{\text{tx}}^{\max})^2\} = n_i/\{\pi(d_{\text{tx}}^{\max})^2(2i-1)\}$. Then, n_i and n can be expressed using n_1 as

$$\begin{aligned} n_i &= (2i-1)n_1, \\ n &= \textstyle\sum_{i=1}^{L} n_i = \sum_{i=1}^{L} (2i-1)n_1 = L^2 \cdot n_1. \end{aligned} \tag{10}$$

By solving $E_2 = E_1$ using (8), (9) and (10), the following relationship can be obtained:

$$T_{c,2} = \frac{2L^2 + 1}{L^2(N_{c,2} + 2)} T_c. \tag{11}$$

Note that as $N_{c,2}$ increases, $T_{c,2}$ becomes zero, which means that if too many cooperators are required, it is too expensive to do CT (so CT task is not used), and our CT model becomes a non-CT model.

In (11), L can be obtained when the network size is fixed, and $N_{c,2}$ is determined based on the amount of the extended range when CT is used. But still, T_c remains unknown. Observe that the number of the tasks that the nodes in A_2 will face is $T_2 = (1 - n_1/n)T_c$, which can be rewritten as $(1 - 1/L^2)T_c$ by using (10). Among these tasks, $T_{c,2}$ tasks (which is (11)) should be selected to cooperate, and, by dividing (11) by $(1 - 1/L^2)T_c$, we get the fraction of the tasks in A_2 that should do CT, which is

$$\frac{T_{c,2}}{T_2} = \frac{2L^2 + 1}{(L^2 - 1)(N_{c,2} + 2)}. \tag{12}$$

This motivates us to use CT with the probability equal to the value in (12). That is, the nodes in A_2 choose to cooperate with the probability in (12) forcing $T_{c,2}$ to satisfy (11). If we define $\text{Prob}_{j,i}$ as the probability of doing CT for a node in A_i when the j-Level CT scheme is used, then (12) is $\text{Prob}_{1,2}$. Again, if $N_{c,2}$ goes to infinity, $\text{Prob}_{1,2}$ in (12) approaches to zero, meaning CT becomes unnecessary. Also, it is not hard to prove that $0 < \text{Prob}_{1,2} < 1$ for $L \geq 2$ and $N_{c,2} \geq 2$. By regulating the CT usage based on $\text{Prob}_{1,2}$, we can achieve E_1=E_2, which is shown to be impossible when non-CT is used [21], [25].

Now, we derive the lifetime extension through CT, which is the increase in the number of tasks that can be performed when CT is used. From the definition in Section 3.3.1, non-CT can perform T_{nc} tasks, and the nodes in A_1 form an energy hole after T_{nc} tasks. However, in the case of the CT, the nodes in A_1 handle $T_{c,2}$ less tasks than non-CT, which means that CT can perform "at least" $T_{c,2}$ more tasks than non-CT until the energy hole is generated. The reason why the increased lifetime (tasks) is not exactly $T_{c,2}$ is because $T_{c,2}$ tasks are "relay" burdens, and relaying a packet consumes more energy than just transmitting a packet (when the node is a source). The increased number of tasks by using CT, which will be denoted by T_x, has the following relationship with T_c:

$$T_c = T_{nc} + T_x. \tag{13}$$

To get the relationship between T_c and T_{nc}, we first express T_x in terms of T_c as follows. The additional T_x tasks are generated by all nodes in the network, and therefore, each node

in A_1 will generate T_x/n tasks on average, and the expected number of tasks generated by nodes in A_1 is $n_1 \cdot T_x/n$ (out of T_x tasks). The rest $T_x - n_1 \cdot T_x/n$ tasks are the expected number of tasks relayed by the nodes in A_1. Therefore, the average energy consumption of the nodes in A_1 to carry out additional T_x tasks can be expressed as

$$\left(T_x - \frac{n_1 \cdot T_x}{n}\right) \cdot E_{\text{relay}} + \frac{n_1 \cdot T_x}{n} E_{\text{TX}}. \tag{14}$$

Note that $T_{c,2}$ is the number of tasks that the nodes in A_1 would have relayed if CT is not used, and therefore, the energy consumption for relaying $T_{c,2}$ tasks should be equal to (14). Solving $T_{c,2} \cdot E_{\text{relay}} = (14)$ gives

$$T_x = \frac{E_{\text{relay}}}{E_{\text{TX}} + (L^2 - 1)E_{\text{relay}}} \times \frac{2L^2 + 1}{(N_{c,2} + 2)} T_c, \tag{15}$$

where we have used (10) and (11).

The relationship between T_c and T_{nc} is obtained by substituting (15) into (13), which is

$$\frac{T_c}{T_{nc}} = \frac{(N_{c,2} + 2)((L^2 - 1)\eta + L^2)}{N_{c,2}L^2 - 1 + (N_{c,2}L^2 - N_{c,2} - 3)\eta}, \tag{16}$$

where $E_{\text{relay}} = (1 + \eta)E_{\text{TX}}$ is used. (16) is the lifetime-extension factor when CT is used, and it is not hard to show that (16) is always larger than one. Therefore, CT always results in performing more tasks than non-CT.

According to Table 2, N_c=2 gives d_{ext}=2.3$d_{\text{tx}}^{\max}$, which means that, with two cooperating nodes, the range extension is more than twice. Therefore, $N_{c,2} = 2$, and if we substitute this value into (16), we get $\{4(\eta + 1)L^2 - 4\eta\}/\{2L^2(1 + \eta) - 5\eta - 1\}$, which is a monotonically decreasing function of L ($\geq$ 2) and converges to two as L goes to infinity (showing the monotonicity is omitted because of its simplicity). This states that the lifetime extension is more than twice regardless of the size of the network (L). Also, for a fixed network size L, (16) is a monotonically increasing function of η for $N_{c,2} = 2$. This states that when the receiving energy portion is increased, it will increase the relaying cost, and the increased relaying cost is more devastating to non-CT than CT because, in the case of non-CT, the nodes close to the sink node have to relay all the packets in the network.

So far, we developed 1-Level CT in which CT is done only by the nodes in A_2. For the nodes in A_k ($k \geq 3$), it is also possible that they can reach the sink node directly with more cooperating nodes, and we derive the CT probabilities, $\text{Prob}_{j,i}$, and the lifetime extension for this case.

Let us consider the 2-Level CT in which the nodes in A_2 and A_3 do CT for $T_{c,2}$ and $T_{c,3}$ tasks, respectively. The relay burden of the nodes in both A_1 and A_2 is reduced by $T_{c,3}$, and the expected number of tasks per node in A_1 when CT is used is $(T_c - T_{c,2} - T_{c,3})/n_1$. Likewise, the expected number of tasks per node in A_2 is $\{T_c(1 - n_1/n) + (N_{c,2} - 1)T_{c,2} - T_{c,3}\}/n_2$. For the nodes in A_3, the actual number of tasks handled by the nodes is $T_c(1 - n_1/n - n_2/n)$, but we need to add the virtual tasks, $(N_{c,3} - 1)T_{c,3}$, to get the energy value. By considering the fact that each node in the network generates T_c/n messages on average, we get E_i's as follows:

$$
\begin{aligned}
E_1 &= \left(\frac{T_c - T_{c,2} - T_{c,3}}{n_1} - \frac{T_c}{n}\right) \cdot E_{\text{relay}} + \frac{T_c}{n} E_{\text{TX}}, \\
E_2 &= \left\{\frac{T_c(1 - n_1/n) + (N_{c,2} - 1)T_{c,2} - T_{c,3}}{n_2} - \frac{T_c}{n}\right\} \cdot E_{\text{relay}} + \frac{T_c}{n} E_{\text{TX}}, \\
E_3 &= \left\{\frac{T_c(1 - n_1/n - n_2/n) + (N_{c,3} - 1)T_{c,3}}{n_3} - \frac{T_c}{n}\right\} \cdot E_{\text{relay}} + \frac{T_c}{n} E_{\text{TX}}.
\end{aligned}
\tag{17}
$$

If we set $E_1 = E_2 = E_3$, the following results are obtained ((10) is used for simplification):

$$
\begin{aligned}
T_{c,2} &= \frac{(2L^2 + 1)N_{c,3} - 4}{L^2(N_{c,3}N_{c,2} + 2N_{c,3} + 4N_{c,2} - 2)} T_c, \\
T_{c,3} &= \frac{4(L^2 + 1)N_{c,2} - 2L^2 + 3}{L^2(N_{c,3}N_{c,2} + 2N_{c,3} + 4N_{c,2} - 2)} T_c.
\end{aligned}
\tag{18}
$$

To obtain the probability of doing CT for the nodes in A_2 and A_3 ($\text{Prob}_{2,2}$ and $\text{Prob}_{2,3}$, respectively), we need to know how many tasks are handled by those nodes. For the nodes in A_3, $(1 - n_1/n - n_2/n)T_c$ tasks ($= T_3$) are handled, and $\text{Prob}_{2,3}$ is obtained by dividing $T_{c,3}$ in (18) by this value. For the nodes in A_2, $(1 - n_1/n)T_c - T_{c,3}$ tasks are handled ($T_{c,3}$ can be

expressed in terms of T_c using (18)), and $\text{Prob}_{2,2}$ can be calculated by dividing $T_{c,2}$ in (18) by this value. The resulting $\text{Prob}_{2,2}$ and $\text{Prob}_{2,3}$ are as follows:

$$\begin{aligned}\text{Prob}_{2,2} &= \frac{(2L^2+1)N_{c,3}-4}{N_{c,3}(N_{c,2}+2)(L^2-1)-(8N_{c,2}+1)},\\ \text{Prob}_{2,3} &= \frac{4(L^2+1)N_{c,2}-2L^2+3}{(L^2-4)(N_{c,3}N_{c,2}+2N_{c,3}+4N_{c,2}-2)}.\end{aligned} \tag{19}$$

Next, the lifetime extension achieved through CT is derived. When 2-Level CT is used, the nodes in A_1 have $T_{c,2}+T_{c,3}$ less "relay" burdens than the case of non-CT. Therefore, to get the lifetime extension for 2-Level CT, we solve $(T_{c,2}+T_{c,3})\cdot E_{\text{relay}} = (14)$ and get T_x as follows ((18) and (10) are used):

$$T_x = \frac{E_{\text{relay}}}{E_{\text{TX}}+(L^2-1)E_{\text{relay}}}T_c \times \frac{(2L^2+1)N_{c,3}+4(L^2+1)N_{c,2}-2L^2-1}{N_{c,3}N_{c,2}+2N_{c,3}+4N_{c,2}-2}. \tag{20}$$

By substituting (20) into (13) and using $E_{\text{relay}} = (1+\eta)E_{\text{TX}}$, we get

$$\frac{T_c}{T_{nc}} = \frac{\{1+(L^2-1)(\eta+1)\}\cdot A}{A+B(\eta+1)}, \tag{21}$$

where $A = N_{c,3}N_{c,2}+2N_{c,3}+4N_{c,2}-2$ and $B = (L^2-1)N_{c,3}N_{c,2}-3N_{c,3}-8N_{c,2}+3$.

From Table 2, we can see that three cooperating nodes ($N_c = 3$) are sufficient to reach the sink node from Level 3, and therefore, $N_{c,3} = 3$. Substituting $N_{c,3} = 3$ and $N_{c,2} = 2$ into (21) gives $9\cdot\{(\eta+1)L^2-\eta\}/\{3L^2(1+\eta)-14\eta-5\}$, which is again a monotonically decreasing function of L (≥ 3) and shows that the lifetime extension is more than three times when 2-Level CT is used.

Instead of providing the results of M-Level CT for all M's, the generalized procedure for getting the probabilities and extended lifetime for M-Level CT is provided in Table 3.

Note that, given a network of size L, it is not hard to see that using the maximum possible M (M-Level CT), which is $L-1$, gives the best lifetime extension. However, there is a limitation in using M=$L-1$ because the nodes located in high levels require many cooperators to reach the sink node directly, and those nodes may not be able to recruit enough cooperators. This is why the generalized M-Level CT ($M \leq L-1$) is required for

Table 3: Generalized procedure for M-Level CT.

1. Generate M+1 linear equations, E_k (k=1,2,...,M+1), using the following formula:

$$\mathrm{E}_k = \left\{ \frac{T_c\left(1 - \sum_{i=1}^{k-1} \frac{n_i}{n}\right) + (N_{c,k} - 1)T_{c,k} - \sum_{i=k+1}^{M+1} T_{c,i}}{n_k} - \frac{T_c}{n} \right\} \cdot E_{\mathrm{relay}} + \frac{T_c}{n} E_{\mathrm{TX}}.$$

   Here, $N_{c,1} = 0$ and $T_{c,1} = 0$.

2. Solve M+1 E_k's for $T_{c,i}$ (i=2,3,...,M+1) using the condition E_1=E_2=...=E_{M+1}.

3. Calculate the number of tasks, $T(i)$ (i=2,...,M+1), handled by the nodes in A_i as follows:

$$T(i) = T_c\left(1 - \sum_{j=1}^{i-1} \frac{n_j}{n}\right) - \sum_{j=i+1}^{M+1} T_{c,j}.$$

4. Using the obtained $T_{c,i}$'s and $T(i)$'s, $\mathrm{Prob}_{M,i}$ is calculated by $\mathrm{Prob}_{M,i} = T_{c,i}/T(i)$.

5. Get T_x in terms of T_c by substituting $T_{c,i}$'s into the following equation:

$$T_x = \frac{\sum_{i=2}^{M+1} T_{c,i}}{\left(1 - \frac{n_1}{n}\right) \cdot E_{\mathrm{relay}} + \frac{n_1}{n} E_{\mathrm{TX}}} E_{\mathrm{relay}}.$$

6. The extended lifetime is obtained by substituting T_x into $T_c = T_{nc} + T_x$ and solving it for T_c.

size L network instead of a single $(L-1)$-Level CT, which gives the best performance. Note that, in order for CT to completely balance the energy consumption of the entire network, $(L-1)$-Level CT is required for the network of size L. Also, when $M < L-1$ and M-Level CT is used, we have $E_1 = E_2 = ... = E_{M+1} > E_{M+2} > E_{M+3} > ... > E_L$[1], which means that the nodes in A_j $(j > M+1)$ do not die earlier than the nodes in A_i $(i \leq M+1)$.

3.3.3 Deployment Method and Cooperative Routing Scheme

So far, we have shown the possibility of overcoming the energy-hole problem under the uniform distribution using CT. As mentioned earlier, we assumed a circular-shaped network used in the non-uniform distribution strategies [21], [25] to claim that balancing energy consumption is possible even with the uniform distribution when our CT method is used. Similar to non-uniform distribution strategies, our protocol also has a deployment method, and, as long as nodes are deployed to meet the deployment criteria, neither the shape of the network nor the distribution of nodes matters, which is discussed below.

Note that the derivation in Section 3.3.2 is based on two conditions introduced in Section 3.2, which are (i) the nodes in adjacent A_i's can communicate with each other, and (ii) the nodes in A_{i-1} equally share the traffic coming from A_i. These two conditions can be achieved through the deployment, and therefore, the deployment method for our CT scheme is to deploy the nodes so that they can satisfy these conditions. One of the simple ways to achieve Condition (ii) is to place nodes regularly in each level. Note that Conditions (i) and (ii) are less strict than the conditions for the non-uniform distribution strategies introduced in [21] and [24]. That is, [21] requires not only a specific number of nodes to be placed in the area but also any node in the network to have q or $q-1$ relay candidates, and [24] assumes high node density and has to place different number of nodes in multiple subregions. Our deployment method does not force a specific number of nodes to be placed in certain areas, which gives more freedom of deployment compared to the non-uniform

[1] From Table 3, we have $E_{M+1} = \{(T_c(1 - \sum_{i=1}^{M} \frac{n_i}{n}) + (N_{c,M+1} - 1)T_{c,M+1})/n_{M+1} - T_c/n\} \cdot E_{\text{relay}} + T_c \cdot E_{\text{TX}}/n$. When the nodes in A_j $(j > M+1)$ only do non-CT, $E_j = \{T_c(1 - \sum_{i=1}^{j-1} \frac{n_i}{n})/n_j - T_c/n\} \cdot E_{\text{relay}} + T_c \cdot E_{\text{TX}}/n$. Using (10), it can be shown that $E_{M+1} > E_{M+2} > E_{M+3} > ... > E_L$.

distribution, not to mention the cost savings (we do not put additional nodes near the sink node to solve the energy-hole problem). Note that the value of $N_{c,i}$ limits how high M can be (M-Level CT). That is, in order to use M-Level CT, nodes in Level i ($2 \leq i \leq M+1$) should have $N_{c,i} - 1$ neighbors that are also in Level i. This means that, given a deployment, the number of neighbors can be a limiting factor of using M-Level CT. In Section 3.4, we will see that, even with 1-Level CT, the lifetime improvement is nontrivial.

Next, we consider the shape and node distribution of the network. Note that, in the derivation, the uniform distribution and the shape of the network are only used in (10) for simplifying n_i, the number of nodes in Level i. This means that $\text{Prob}_{j,i}$ can be derived for arbitrary network shapes and various node distributions as long as they conform to the proposed deployment strategy. For example, we can derive $\text{Prob}_{j,i}$ and T_c for the strip network having the same number of nodes in each level by setting $n_1 = n_2 = ... = n_L$ and going through the generalized procedure in Table 3.

We now discuss our routing protocol. The cooperative routing protocol that we have devised based on our analysis can be implemented on top of any existing cross-layer framework designed to use range-extension CT, and such a cross-layer framework is well introduced in [47], which enables a node to gather cooperators and do CT. Also, as in [47], when a source node wants to send a data, it first establishes a non-CT primary route to the destination, which is identical to the operation of conventional non-CT routing protocols. This makes it possible for a node to do non-CT (i.e., forwarding data to the next-hop node without doing CT) when the node decides not to do CT based on the probability $\text{Prob}_{M,i}$.

To use M-Level CT, one must fix the network topology, which fixes L. Then, $\text{Prob}_{M,i}$ can be obtained since M and $N_{c,i}$ can be decided in advance. Each node can save $\text{Prob}_{M,i}$ and be deployed. Once deployed, nodes need to figure out which level they are in (the number of hops they are away from the sink node), which can be done when the network is initialized [55]. After the initialization phase, each node in A_i ($2 \leq i \leq M+1$), whenever it has a task to perform (transmitting or relaying a message), chooses to do CT with the

probability $\text{Prob}_{M,i}$, where i is the level the node is in. When a node decides to do CT, it cooperates with neighboring cooperators that are in the same level. It is important to note that $\text{Prob}_{M,i}$ is used by the node that has a packet to send, and, when the node decides to do CT, the cooperators chosen by the node should always cooperate (help). If the node decides not to do CT, the node performs the conventional non-CT routing protocol (i.e., it routes the packet to the next-hop node on the primary route)[2].

We refer to the cooperative routing scheme (along with the deployment method) introduced in this section as "proactive cooperative transmission (PROTECT)."

3.4 Simulation

In this section, we verify our analysis of Section 3.3.2 and evaluate the lifetime performances of PROTECT under different conditions through network simulations using MATLAB. In simulation, a node selects a route based on its routing decision, and the residual energy of a node is reduced according to the energy model (below) when a node receives or transmits a packet. In simulation, we maintain the network and physical-layer models in Sections 3.2 and 2.4.2. The maximum SISO communication range of a node is set to 40m ($d_{\text{tx}}^{\max} = 40\text{m}$), and $N_{c,2}$=2 and $N_{c,3}$=3 are used following the values in Table 2. E_{TX}=0.092mJ and E_{RX}=0.277mJ are used ($\eta \approx 3$) following the energy model in [19] and assuming 256 bytes of data, and we choose the initial energy of each node to be 100mJ. Note that we choose the initial energy suitably so that simulation can be done in timely manner[3].

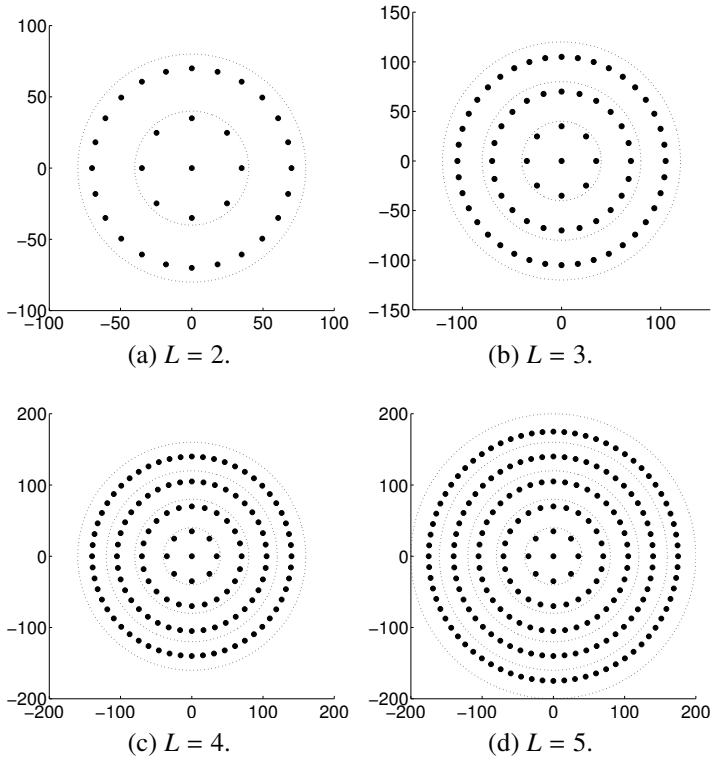
3.4.1 Verification of Theoretical Values

Here, we check the validity of the probabilities ($\text{Prob}_{j,i}$) and the extended lifetime derived in Section 3.3 through our network simulations. We first consider the circular-shaped network

[2]Note that the nodes in A_1 and A_j ($j > M+1$) perform only the conventional non-CT routing protocol.

[3]The value of the initial energy does not play a role in our analysis. We've simulated PROTECT using different values of the initial energy (50mJ, 1J, etc.), and we were able to verify our analysis for those values as well. Therefore, the results with different initial energies are omitted.

with the uniform distribution, which is used in Section 3.3.2. L=2, 3, 4, and 5 are considered, and the number of nodes are 32, 72, 128, and 200 respectively, which corresponds to n_1=8 in (10). Figure 7 shows the network topologies with different sizes (different L's) that we have used (the dotted circle indicates each level), where nodes are deployed to satisfy our deployment method.



(a) L = 2. (b) L = 3.

(c) L = 4. (d) L = 5.

Figure 7: Four circular-shaped network topologies (L=2, 3, 4, and 5) with $d_{\text{tx}}^{\max}$ = 40m. The dotted circles indicate C_i's.

As already mentioned in Section 3.2, shortest-hop routing is used for both non-CT and CT (primary routing in the case of CT). The simulation is done till no packets can reach the sink node (no connectivity to the sink node). We evaluate how the connectivity changes as the number of tasks increases. The sources are randomly selected, and we repeat the tests 25 times and average the results.

Figure 8 shows the number of connected nodes versus the number of tasks performed

when 1-Level PROTECT with $\text{Prob}_{1,2}$=0.59 (obtained from (12)) is used for the L=3 network. Note that T_c and T_{nc} are the number of tasks that can be performed before the network is "completely" disconnected for CT and non-CT, respectively. Since Figure 8 is showing the averaged values, the point where the connectivity goes below one is regarded as T_c and T_{nc}. These points are 2375 for non-CT and 5545 for PROTECT, and the lifetime-extension factor is 2.34, which is close to the theoretical value 2.36 obtained from (16). The slight mismatch between the theoretical extension value and the actual extension value is inevitable because (i) when source nodes are selected randomly, there is no guarantee that each node will evenly generate and receive the exact expected number of tasks used in the derivation of Section 3.3 and (ii) in the ideal (theoretical) situation, all nodes should die at the same time, but in reality, some nodes should die first, and this early death causes a change in the network shape, and thereby a deviation from the derivation.

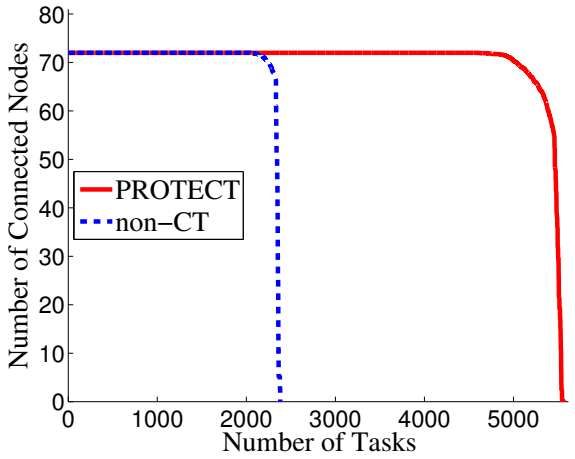


Figure 8: The simulation result for 1-Level PROTECT. L=3 circular-shaped network.

In addition to T_c and T_{nc}, we provide the values when the first node dies, which is a widely used definition of the network lifetime. In order not to be confused with the terms, this definition of lifetime will be referred to as "first death lifetime (FDL)," and T_c and T_{nc} are "last death lifetimes (LDLs)." FDL is measured by getting the point when the connectivity falls below 'maximum connectivity−1'. The simulation results of 1-Level and 2-Level PROTECT for all L's (L=2, 3, 4, and 5) are summarized in Table 4, which clearly

Table 4: The simulation results for 1-Level and 2-Level PROTECT.

1-Level PROTECT				
	$L = 2$	$L = 3$	$L = 4$	$L = 5$
$\text{Prob}_{1,2}$	0.75	0.59	0.55	0.53
LDL Extension - Theoretical	3.25	2.36	2.18	2.11
LDL Extension - Simulation (T_c/T_{nc})	3.21 (8669/2698)	2.34 (5545/2375)	2.17 (4953/2287)	2.10 (4702/2240)
FDL Extension - Simulation	3.28	2.26	2.06	2.01
2-Level PROTECT				
	$L = 2$	$L = 3$	$L = 4$	$L = 5$
$\text{Prob}_{2,2}$ / $\text{Prob}_{2,3}$	N/A	0.67 / 0.72	0.58 / 0.50	0.55 / 0.43
LDL Extension - Theoretical	N/A	4.87	3.79	3.45
LDL Extension - Simulation (T_c/T_{nc})	N/A	4.81 (11430/2375)	3.75 (8572/2287)	3.44 (7704/2240)
FDL Extension - Simulation	N/A	4.83	3.64	3.24

shows the similarities between theoretical and simulation values[4]. Note that we don't have theoretical FDL extension values, and only the simulated values are provided in Table 4.

To see if the energy consumption is actually balanced for PROTECT, we provide Figure 9, which shows the average residual energy per node (in Joules) for each level right after one node is dead when non-CT and 2-Level PROTECT are used for the L=3 network[5]. As can be seen from the figure, PROTECT well balances the energy consumption compared to non-CT. Figure 9 also shows that non-CT leaves a huge amount of energy unused explaining why PROTECT is so effective.

So far, we have verified the derivation in Section 3.3.2. In order to verify the claim in Section 3.3.3 that our CT scheme can be applied to other shapes of the network, we

[4]The connectivity vs. task graphs for the simulations in Table 4 are omitted except for the case of 1-Level PROTECT with L=3 (in Figure 8) because their decreasing trends and shapes are almost identical to those of Figure 8.

[5]The standard deviations (25 samples) of Level 1, 2, and 3 are 0.00004, 0.0007, and 0.00007, respectively for non-CT, and 0.00004, 0.0001, and 0.0004, respectively for 2-Level PROTECT.

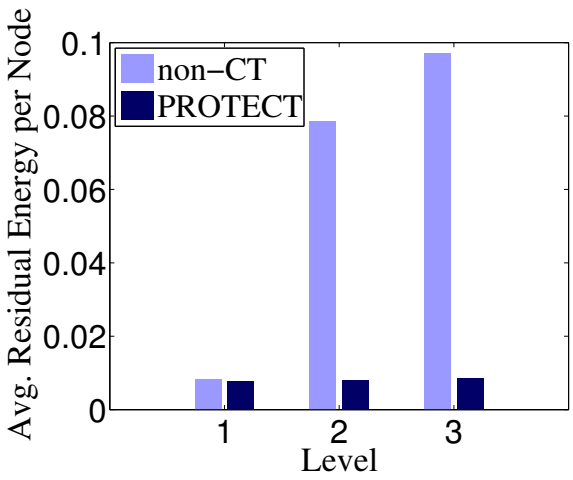


Figure 9: The average residual energy (in Joules) per node for each level when 2-Level PROTECT is used for L=3 circular-shaped network.

consider a strip network having the same number of nodes in each level. We consider L=2 and L=3 networks with six nodes in each level, each of which can talk to each other, and we consider 1-Level PROTECT for L=2 and 2-Level PROTECT for L=3. For these cases, it can be derived (by using the generalized procedure in Table 3) that $\text{Prob}_{1,2}$ =1/2 and $T_c/T_{nc} = (2\eta + 4)/(\eta + 3) = 1.67$ for L=2 network, $\text{Prob}_{2,2}$ =1/3, $\text{Prob}_{2,3}$=1/2, and $T_c/T_{nc} = (2\eta + 3)/(\eta + 2) = 1.8$ for L=3 network. In the simulation, the sources are randomly selected, and we repeat the tests 25 times and average the results. The simulation gives the LDL extension values of 1.66 for L=2 and 1.79 for L=3, which match well with the theoretical values 1.67 and 1.8, respectively.

3.4.2 Effect of Suboptimal Probabilities

In this subsection, we see the effect when the probabilities different from the optimal (theoretical) value are used. We consider the circular-shaped networks used in Section 3.4.1. Figure 10 shows the result for 1-Level PROTECT when L=3. The probability of 0.59 is the theoretical value ($\text{Prob}_{1,2} = 0.59$), and the simulation results when probabilities are 0.49, 0.54, 0.59, 0.64, and 0.69 are provided along with the non-CT case. When the probabilities are less than the theoretical value, i.e., 0.49 and 0.54, the graph clearly shows that their

performances are worse than the performance of the optimal probability 0.59. Interesting results are obtained when the probabilities are larger than 0.59, which gives longer LDLs than the case of 0.59. Also, it can be observed that these probabilities have a transition point where the steep slope changes to a mild (flat) one. The main reason why the probabilities larger than 0.59 can perform more tasks is because of the fact that the task *originating* from higher level will consume more energy to complete the task. That is, the energy consumed when the task is *originating* from A_1 is E_{TX}, whereas, in the case of the task generated in A_2, $2E_{\mathrm{TX}}+E_{\mathrm{RX}}$ is required to finish the task. Therefore, maximizing the tasks originating from A_1 can maximize the number of overall performed tasks. In the case where the probability is larger than 0.59, the nodes in A_2 die early, and this lets more tasks to be generated by the nodes in A_1. The transition points in Figure 10 indicate the death of all nodes in A_2, and since the node in A_1 only needs to transmit (using E_{TX}) its own data instead of relaying (using $E_{\mathrm{TX}}+E_{\mathrm{RX}}$) after the early death of the nodes in A_2, more tasks can be performed than the case where this early death does not occur (the case of using the optimal probability).

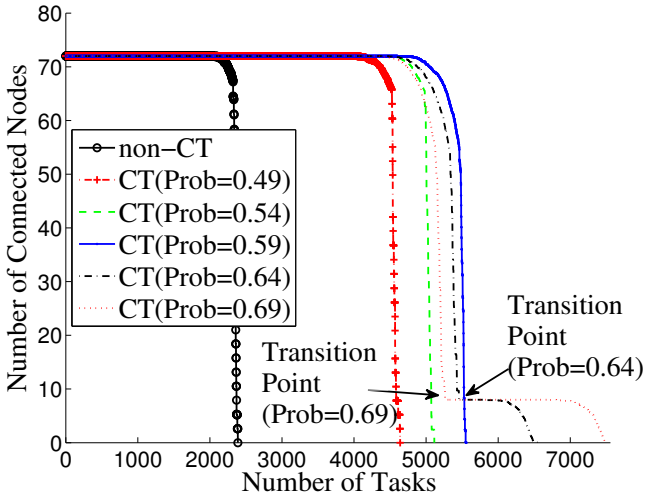


Figure 10: The effect of different probabilities. 1-Level PROTECT with different $\mathrm{Prob}_{1,2}$. L=3 circular-shaped network.

The FDLs for probabilities 0.49, 0.54, 0.59, 0.64, and 0.69 are 4285, 4716, 4952, 4767, and 4641, respectively. Note that the longest FDL is obtained when the optimal probability 0.59 is used, and this indicates that using the optimal probability not only prevents the

early death of nodes but also well balances the energy compared to the other cases. To see the energy balancing for different probabilities more clearly, we provide Figure 11, which shows the average residual energy per node in Level 1 and 2 right after one node is dead. The result shows that when the probability is less than the optimal value, 0.59, the nodes in A_1 consume more energy (have less residual energy) than the nodes in A_2, whereas, in the opposite case, the nodes in A_1 consume less energy.

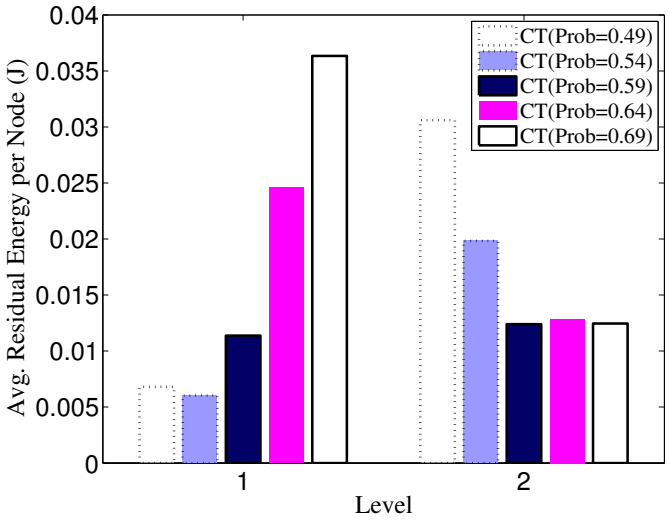


Figure 11: The average residual energy per node in Level 1 and 2 right after one node is dead with different probabilities for 1-Level PROTECT with $L = 3$. The results are obtained from the same simulation of Figure 10.

In the case of 2-Level PROTECT, Figure 12 shows the effect of the different probabilities when $\text{Prob}_{2,2}$ is fixed to the optimal value and $\text{Prob}_{2,3}$ is varied. Here, the L=4 network is used, and the optimal values for $\text{Prob}_{2,2}$ and $\text{Prob}_{2,3}$ are 0.58 and 0.5, respectively. Similar to the case of 1-Level PROTECT, the transition points exist for the probabilities higher than the theoretical value, 0.5, and the probabilities less than 0.5 underperform. The transition points indicate the death of all nodes in A_3 because of the excessive usage of CT. Note that the early death of outer nodes can be devastating. For example, in Figure 12, for the probability of 0.6, about 75% of connectivity is lost when 8000 tasks are performed, whereas, for the probability of 0.5 (optimal), the connectivity is nearly full. From Figure

12, we can see that the FDL value is the highest when the optimal probability, 0.5, is used. Although the results are not provided here, the same trend is observed when $Prob_{2,3}$ is fixed to 0.5 and $Prob_{2,2}$ is varied.

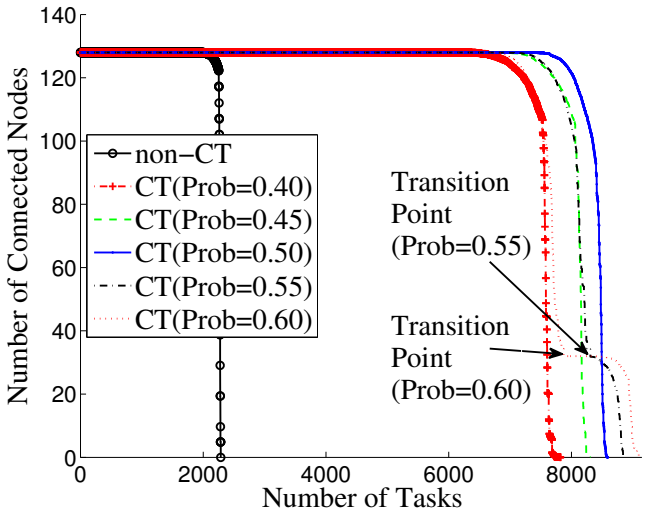


Figure 12: The effect of different probabilities. 2-Level PROTECT with different $Prob_{2,3}$. $Prob_{2,2}$=0.58, and circular-shaped networks with L=4 is used.

To summarize, the case where the probability is larger than the theoretical value can perform more tasks, but the balancing of energy fails. This unbalanced energy leads to the creation of an *outer* energy hole and disconnects all nodes far away from the sink node early. Therefore, unless the data near the sink node is more valuable than the data far away from the sink node, using probabilities higher than the optimal value is not so desirable, and using the optimal probability can give the best performance.

3.4.3 Non-Uniform Deployment

So far, we have focused on verifying our expected lifetime extensions and theoretical CT probabilities. In this section, we discuss how the non-uniform deployment affects the performance of PROTECT, and we show that PROTECT can always help, unless quite a lot of nodes are added.

We focus on a two-hop circular-shaped network (L=2), which is simple but can sufficiently address the benefits of using PROTECT. Note that the non-uniform distribution

strategy in [21] assumes that all nodes are transmitting their own data. However, it is also possible that some nodes that are placed in Level 1 do not have to generate their own data because their main role is to reduce the relay burden, and reporting their data does not contribute to the sensing coverage in Level 1. In order to capture this fact, we define n_1 as the number of nodes in Level 1 that generate packets of their own, and n_1^{ADD} as the number of nodes in Level 1 that are placed strategically to reduce the relay burden of n_1 nodes. We consider two cases in this section where (i) the additional nodes transmit their data ($n_1^{\mathrm{ADD}} = 0$, and we vary n_1) and (ii) the additional nodes only relay ($n_1^{\mathrm{ADD}} > 0$, and we vary n_1^{ADD}).

Considering the fact that n_1+n_1^{ADD} nodes in Level 1 are relaying, and each node in Level 1 generates $n_1 \cdot T_c/\{n(n_1 + n_1^{\mathrm{ADD}})\}$ tasks on average, we get

$$E_1 = \left(\frac{T_c - T_{c,2}}{n_1 + n_1^{\mathrm{ADD}}} - \frac{n_1 \cdot T_c}{n(n_1 + n_1^{\mathrm{ADD}})}\right) \cdot E_{\mathrm{relay}} + \frac{n_1 \cdot T_c}{n(n_1 + n_1^{\mathrm{ADD}})} E_{\mathrm{TX}}, \tag{22}$$

where n is n_1+n_2, not n_1+n_1^{ADD}+n_2. Since the energy consumptions of the nodes in Level 2 do not change, E_2 is still (9), and we use the procedure in Table 3 to get $\mathrm{Prob}_{1,2}$ and the lifetime-extension factor for 1-Level PROTECT, which are

$$\begin{aligned} \mathrm{Prob}_{1,2} &= \frac{n_2(1+\eta) - n_1^{\mathrm{ADD}}}{(n + n_1^{\mathrm{ADD}})(1+\eta)}, \\ \frac{T_c}{T_{nc}} &= \frac{n_1^{\mathrm{ADD}}\{n_2(\eta+1) + n_1\} + n(n + n_2\eta)}{n_1^{\mathrm{ADD}}\{n_2(\eta+2) + n_1\} + n_1(n_1 + 2n_2 + n_2\eta)}. \end{aligned} \tag{23}$$

Note that $n_1^{\mathrm{ADD}} = 0$ is the case of [21], where all nodes in Level 1 are sending their own sensed data, and, in this case, $\mathrm{Prob}_{1,2}$ is always positive, which means that CT is always required to balance the energy consumption. On the other hand, the case where $n_1^{\mathrm{ADD}} > 0$, $\mathrm{Prob}_{1,2}$ is zero when $n_1^{\mathrm{ADD}} = n_2(1+\eta)$, which means that $n_2(1+\eta)$ "redundant" nodes are required in Level 1 to balance the energy consumption without CT[6]. When $n_1^{\mathrm{ADD}} > n_2(1+\eta)$,

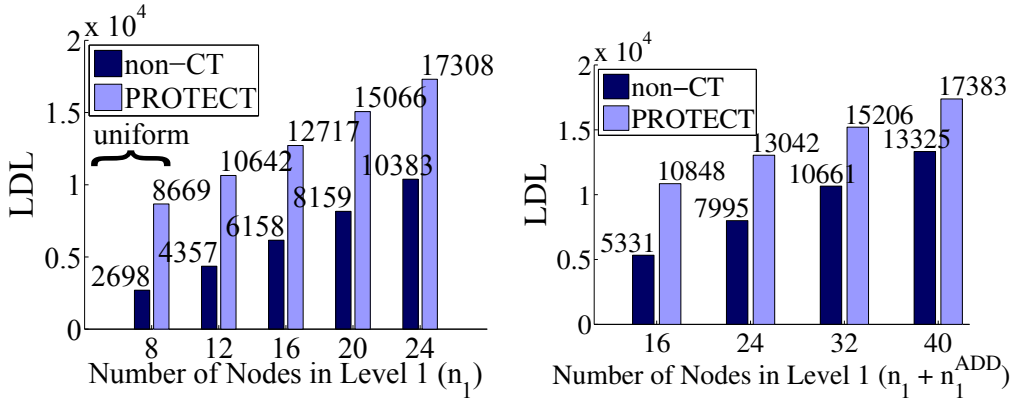
[6]The following simple example shows that $n_1^{\mathrm{ADD}} = n_2(1+\eta)$ actually makes sense. Consider the case where n_1=1, n_2=1, E_{TX}=1, and E_{RX}=3 and assume that each node has the energy of 4. The single node in A_2 can send four packets because E_{TX}=1, and, in order for the single node in A_1 to send four packets, additional nodes are required in A_1 to relay the packets coming from A_2. The relaying energy cost is 4, and, since a node has enough energy to relay one packet only, n_1^{ADD}=4, which is $n_2(1+\eta)$.

$\text{Prob}_{1,2}$ is negative (which is impossible), and this indicates that CT is unnecessary because the redundant nodes placed in A_1 makes $E_1 < E_2$.

As for the lifetime-extension factor in (23), in the case where n_1^{ADD}=0, the extension factor converges to one as n_1 goes to infinity (also $\text{Prob}_{1,2}$ becomes zero), which indicates that if we can increase the number of nodes in Level 1 to be large enough, using CT is no longer powerful because the non-uniform distribution strategy alone can successfully deal with the energy-hole problem. Note that to actually reach the point where CT's lifetime extension becomes negligible, an excessive amount of nodes are required in Level 1. For example, with n_2=24 and η used in Section 3.4.1, n_1=100 gives the lifetime-extension factor of 1.1, and more than 424 nodes are required in Level 1 (n_1 = 424) to get the lifetime-extension factor less than 1.01. In the case of $n_1^{\text{ADD}} > 0$, having $n_1^{\text{ADD}} = n_2(1 + \eta)$ nodes gives T_c=T_{nc}.

We now provide the network simulation results for the circular-shaped network (L=2) with the non-uniform deployment. For reference, the two-hop circular-shaped network (L=2) with uniform deployment in Figure 7a has n_1=8 and n_2=24, according to (10). Therefore, in the simulation, we increase $n_1 + n_1^{\text{ADD}}$ and obtain the performances of PROTECT and non-CT while n_2 is fixed. We provide two cases: (i) $n_1^{\text{ADD}} = 0$ and (ii) $n_1^{\text{ADD}} > 0$. When $n_1^{\text{ADD}} > 0$, we fix n_1 to 8 and increase n_1^{ADD}. The test conditions are exactly the same as the tests on the circular-shaped networks that we used in Section 3.4.1, and the only difference is the number of nodes in Level 1. Figure 13 shows the LDL values of the non-uniform deployment obtained from network simulations. In Figure 13a, we also put the result of the uniform deployment case as a reference. As can be seen from the figures, PROTECT can provide better lifetime performance than non-CT under the non-uniform deployment. More importantly, PROTECT under the uniform deployment can have better performance than non-CT using the non-uniform distribution strategy, which requires additional nodes. For example, when $n_1^{\text{ADD}} = 0$, PROTECT with n_1=8 has better lifetime performance than the (non-CT) non-uniform case with n_1=20 (8669 vs. 8159). In other words, the non-uniform

distribution strategy requires more than 12 additional nodes (37.5% more nodes than PROTECT) to perform as well as PROTECT. When $n_1^{\mathrm{ADD}} > 0$, more than 16 additional nodes are required for the non-uniform distribution strategy to match with PROTECT under the uniform deployment (8669 vs. 7995).



(a) $n_1^{\mathrm{ADD}} = 0$. (b) $n_1 = 8$ and $n_1^{\mathrm{ADD}} > 0$.

Figure 13: The LDL results of the non-uniform deployment. Circular-shaped network with L=2 and n_2=24.

We summarize the LDL extension factors and FDL values in Table 5. From the table, we can see that the theoretical values obtained from (23) match well with the simulation values. We can observe that, as the number of nodes in A_1 increases, the LDL extension factor decreases because the non-uniform deployment can reduce the relay burden of the nodes in A_1.

$T_c/T_{nc} > 1$ until the non-uniform deployment can completely balance the energy consumption without CT, which requires a lot of additional nodes (for $n_1^{\mathrm{ADD}} > 0$, about 96 nodes are required), and this means that PROTECT can always be beneficial even under the non-uniform deployment unless the non-uniform distribution strategy can completely solve the energy-hole problem.

3.4.4 Other Considerations

So far, we have focused on verifying our analysis through simulations. In this section, we discuss the cases that are out of the scope of the analysis in Section 3.3.

Table 5: The LDL extension factors and FDLs for L=2 circular-shaped network under the non-uniform deployment.

$n_2 = 24$.	**Non-uniform**: $n_1^{\mathrm{ADD}} = 0$				**Non-uniform**: $n_1^{\mathrm{ADD}} > 0$			
	n_1				$n_1 + n_1^{\mathrm{ADD}}$ ($n_1 = 8$)			
Total nodes in Level 1 in Level 1	12	16	20	24	16	24	32	40
LDL Extension - Simulation	2.44	2.06	1.85	1.67	2.04	1.63	1.43	1.31
LDL Extension - Theoretical	2.45	2.05	1.82	1.67	2.03	1.63	1.42	1.30
non-CT FDL	3999	5295	7024	9448	4887	6409	9501	10454
PROTECT FDL	9993	11213	13102	15894	9970	10593	13550	13423

As we have introduced in Section 2.1, to extend the lifetime of multi-hop networks using network-layer approaches, many authors have proposed energy-aware routing protocols. Since the burden of the nodes one hop away from the sink node cannot be reduced even with the energy-aware routing (as we have addressed in Section 2.2), we can expect that the energy-aware routing cannot give the lifetime advantage in LDL, and this is the reason why non-CT shortest-hop routing was used for our analysis. However, energy-aware routing schemes can have a clear advantage in FDL because they can detect the node with low residual energy and avoid it from being used. Therefore, in this section, we observe the FDL performance of the energy-aware routing and compare it with that of PROTECT. Here, we add an energy-aware routing protocol, CMAX (introduced in Section 2.1.3), in our simulation. The reason why we have chosen CMAX is that, as we have mentioned in Section 2.1, it has been shown to be one of the most effective online energy-aware routing protocols.

The simulation results of the circular-shaped networks with the uniform distribution used in Section 3.4.1 are summarized in Table 6, which shows FDL and LDL values. As can be seen from the table, there are no notable differences in LDL values between the shortest-hop routing and CMAX, whereas, in the case of FDL, CMAX performs notably better than

Table 6: The FDL and LDL results of the circular-shaped networks with the uniform distribution including CMAX.

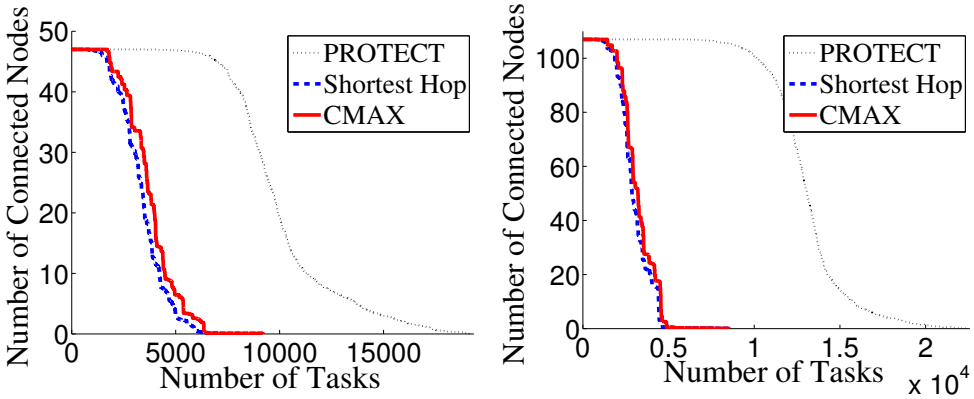
L and n	L=3 (n=72)		L=4 (n=128)		L=5 (n=200)	
Lifetime	FDL	LDL	FDL	LDL	FDL	LDL
CMAX	2342	2376	2250	2289	2218	2251
Shortest-Hop	2195	2375	2134	2287	2085	2240
1-Level PROTECT	4957	5545	4399	4953	4188	4702
2-Level PROTECT	10590	11430	7754	8572	6784	7704

the shortest-hop routing because the shortest-hop routing cannot avoid the node having a low residual energy from being selected. Note that PROTECT also does not use the residual energy information and relies on the shortest-hop routing, and the results clearly show that, even without the knowledge of the residual energy, PROTECT can outperform the non-CT energy-aware routing protocol.

Now, we consider the case where nodes are deployed randomly and uniformly. The random deployment is one of the cases where our deployment criteria discussed in Section 3.3.3 cannot be met, and therefore, it is worth observing how our proposed routing method is affected by the random deployment.

We consider L=2 circular-shaped networks (a radius of 80m) with n=48 and L=3 circular-shaped networks (a radius of 120m) with n=108 where 1-Level PROTECT and 2-Level PROTECT are used, respectively. For PROTECT, we use $\text{Prob}_{j,i}$'s of the uniform deployment case. Also, the network topology is changed in each trial, and we increase the number of trials to 50 and average the results. Figure 14 shows the connectivity vs. task graphs for PROTECT, shortest-hop routing, and CMAX when nodes are deployed randomly. It can be observed from Figure 14 that both non-CT and CT no longer have a sharp decrease in the connectivity. This is because we are using the random deployment, and we change the topology in each trial. That is, when the topology changes in a random deployment, it changes the network lifetime greatly depending on how many nodes are in A_1 and whether

there exists a direct link between a node in A_1 and a node in A_2. Also, in a random deployment, there is a high possibility that there are multiple bottleneck nodes in not only A_1 but also A_2, and the death of one of these nodes can cause a major loss of connectivity. These large variances in both lifetime and connectivity are the main reason why we get distorted graphs for both non-CT and CT in Figure 14, which are showing the averaged results. Note that non-CT and CT share similar decreasing trends, and CT still gives a notable lifetime advantage.

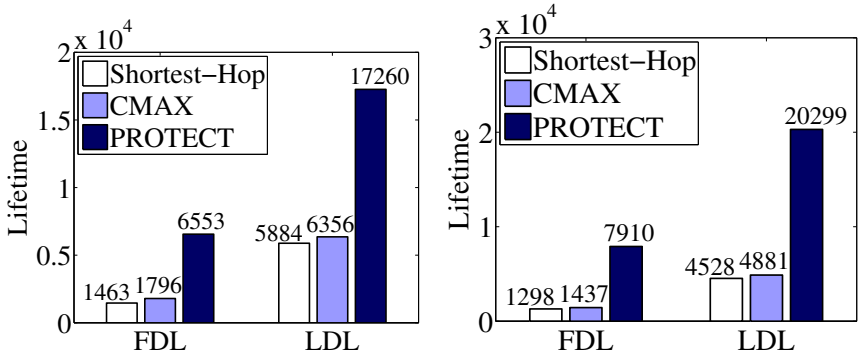


(a) L=2 and n=48. M=1.

(b) L=3 and n=108. M=2.

Figure 14: Connectivity vs. task graphs for the random deployment.

Figure 15 clarifies the advantage of using range-extension CT by showing the FDL and LDL values of the random deployment. It can be observed that the extension factors of FDL and LDL are more two when 1-Level PROTECT is used for L=2, and, in the case of L=3 with 2-Level PROTECT, the extension factors of FDL and LDL are more than four. Even in the case of the random deployment, the nodes close to the sink node still suffer from high relaying cost, and PROTECT, by reducing the relaying burden of those nodes, can largely improve the network lifetime. However, unless the random deployment happens to satisfy the deployment criteria in Section 3.3.3, PROTECT cannot guarantee the balanced energy consumption under the random deployment.



(a) $L = 2$ and $n = 48$. (b) $L = 3$ and $n = 108$.

Figure 15: The LDL and FDL values of the random deployment.

3.5 Discussion

3.5.1 Protocol Overhead of PROTECT

Any CT protocol requires a node to share its data with cooperator(s), and, when a node decides to do CT, the node may or may not need to explicitly select cooperator(s) for each CT transmission depending on how the underlying protocol is designed. For example, when two cooperating nodes are required for CT, the two nodes can agree to form a pair during the network initialization phase (or using hard coded pairs and a careful deployment) so that if one does CT, the other can help, and, in this case, no explicit cooperator selection is necessary for each CT transmission. Note that the data sharing for CT can be done when the first nodes (that has a packet to send) transmits, which is the case used in [12], and during this phase, the second node can know its participation in CT. Since the data sharing of CT and a node's helping the other node's CT (i.e., transmitting other node's packet) are included in our analysis as a virtual task, when no explicit cooperator selection is necessary for each CT transmission, our analysis should hold.

There can also be the cases where a node should select different cooperator(s) for each CT transmission, and there can be many ways to do this. One example is in [47], which uses an existing control packet (modified RTS) to select cooperators[7]. One can also think about

[7]Note that [47] also addresses how multiple CT flows can be handled using their cross-layer framework, which is not considered in this dissertation.

putting additional header field that indicates cooperators when the data packet is transmitted in the first time slot. In any case, when explicit cooperator selection is necessary for each transmission, additional bytes in either a control packet or a data packet (these are the overhead for using CT that "uniquely" exists when CT is used and is "not" necessary for non-CT) may be required when doing CT, which leads to the additional energy consumption for PROTECT. To consider the additional energy consumption theoretically, let us assume that X additional bits are required for each packet transmission of PROTECT so that these bits affect the energy consumption for each transmission and reception. Note that this is a very crude (energy-inefficient) approach of utilizing PROTECT because the nodes in A_1 does not do CT at all in PROTECT, but those nodes are still assumed to be consuming energy for the additional bits. If we denote by α, the ratio of (additional) X bits to the length of the entire non-CT data packet in bits, then, when 1-Level PROTECT is used (we only consider 1-Level PROTECT in this section), we can see that the probability ($\text{Prob}_{1,2}$) derivation does not change because we are now solving $(1 + \alpha) \times (8) = (1 + \alpha) \times (9)$ instead of (8) = (9) used in Section 3.3.2. However, the lifetime-extension factor changes, and this value can be obtained by (i) considering the total energy consumption for a node in A_1 when 1-Level PROTECT and non-CT are used, which are $(1 + \alpha) \times (8)$ and $T_{nc}\{(1/n_1 - 1/n) \cdot E_{\text{relay}} + E_{\text{TX}}/n\}$ respectively, and (ii) setting these two energy consumption values to be equal (because the total energy consumption is bounded by the initial energy regardless of using PROTECT or non-CT). The resulting lifetime-extension factor can be shown to be $(16)/(1 + \alpha)$. Since the method we assumed is very crude (energy-inefficient), we can claim that the lifetime obtained using this method is the lower bound for PROTECT, which leads to the following condition:

$$\frac{(16)}{(1+\alpha)} \leq \frac{T_c}{T_{nc}} \leq (16). \tag{24}$$

In Figure 16, we plot the lower (dash-dot line) and upper (solid line) bounds of T_c/T_{nc} using the same test conditions used in Section 3.4.1 for $L = 2$ network. For selected α's (0.5, 1, 1.5, 2, and 2.5), we also performed network simulations and verified the validity, which is

also shown in Figure 16. As can be seen from Figure 16, the overhead negatively affects the performance of PROTECT. However, according to Figure 16, in order for PROTECT to become useless because of the overhead, it requires unrealistically large overhead. That is, T_c/T_{nc} goes below one when α is around 2.25, which means that the overhead is more than twice the size of the entire non-CT packet. If we design the overhead to be only a few bytes, and if we assume that the entire non-CT packet is more than 100 bytes, then α should be less than 0.1, and according to Figure 16, the lifetime extension is more than three times when 1-Level PROTECT is used.

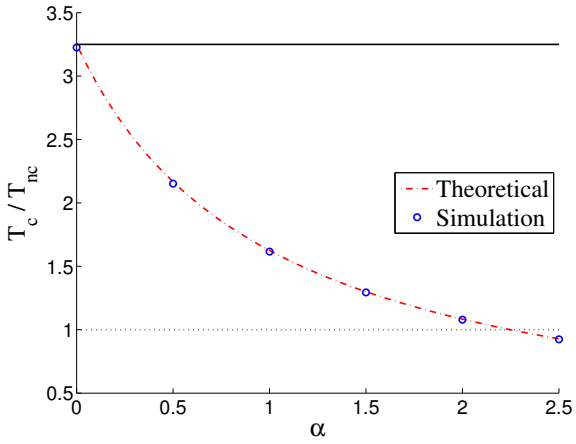


Figure 16: Lifetime-extension factor. $L = 2$, $M = 1$, and the circular-shaped network used in Section 3.4.1.

As we have mentioned earlier, the exact value of the overhead for CT (that uniquely exists for CT only) depends on how one designs the overall CT framework, which usually includes the medium access control (MAC) layer, and whether the CT overhead can be made to be negligible or not is out of scope of this dissertation.

3.5.2 Physical-Layer Consideration

In Sections 2.4.2 and 3.2, we've assumed a certain physical-layer model for our analysis, which affects the number of cooperating nodes required to directly communicate with the sink node successfully. If the channel is different from what we have used in this dissertation, and if the diversity and array gains are not large enough for that channel to use

PROTECT, one may need to use more cooperating nodes to deal with the situation, which affects the lifetime performance of PROTECT. We can prove that (16) is a monotonically increasing function of η and a decreasing function of L. By setting $\eta = 0$ and letting L go to infinity, we get $(N_{c,2} + 2)/N_{c,2}$, which is the minimum lifetime-extension factor when 1-Level PROTECT is used. Even with $N_{c,2}$=4 (twice the value that we have used), the minimum lifetime-extension factor is 1.5, and we note that this is a very pessimistic minimum because $\eta = 0$ and $L = \infty$ are not possible. For the energy and network models we have used in Section 3.4 with L=2, the lifetime-extension factors when 1-Level PROTECT is used are 2.24 and 1.86 for $N_{c,2}$=3 and 4, respectively, which are nontrivial.

3.5.3 Implementation Considerations of PROTECT

There are two efficient techniques that can provide a more accurate operation of PROTECT, which are discussed in this section.

In order for a node to meet $\text{Prob}_{j,i}$, a node can generate a random number between zero and one and decide to do CT when the random number is less than or equal to $\text{Prob}_{j,i}$, which can be done every time the node receives a packet or predetermined before the arrival of a packet. A more simple and effective method is for a node to have a deterministic program of CT and non-CT usage, such that the moving average over any window of sufficient length can approximate the desired probability. This can be achieved by choosing to do CT so that "cumulative" probability of doing CT is close to $\text{Prob}_{j,i}$. That is, when the probability $\text{Prob}_{j,i}$ is 0.74, the node should do CT at the first arrival of the packet because one (doing CT out of one arrival) is close to 0.74 than zero (not doing CT out of one arrival), and the second arrival, it should not cooperate because doing CT gives the cumulative probability of one (two out of two arrivals), whereas not doing CT gives the cumulative probability of 0.5 (one out of two arrivals), and 0.5 is close to 0.74 than one. Given $\text{Prob}_{j,i}$, this method (which can be easily implemented) can predetermine whether to do CT or not, and it is not hard to see that this method makes the portion of doing CT stay close to $\text{Prob}_{j,i}$ all the time and converges faster to $\text{Prob}_{j,i}$ than the case of generating random numbers.

When a node has more than enough available cooperators, the node has to pick cooperators. One of the ways to pick cooperators is to select nodes randomly among available cooperators, but a more efficient way is to select them based on the usage history. That is, the node can remember its usage of available cooperators, and, based on this usage information, neighbors with the least usage can be selected. By using this method, each cooperator can share similar amount of CT flows, and no node has to die early because of the uneven distribution of the CT burden.

3.6 Summary

In this chapter, we have presented an analytical model for avoiding the energy hole using range-extension CT. Through analysis, we have calculated the optimal probability of doing CT, which indicates how often CT should be performed, and, based on the analysis, we have provided a unique CT scheme that requires only the level information obtained during the network initialization period to avoid the energy hole. We also derived how much lifetime extension our CT method can get, and both the optimal probability and extended lifetime values were verified by network simulations. By using PROTECT, it is possible to balance the energy consumption between the nodes in different levels and avoid the energy hole even with the uniform distribution of nodes. Also, because of the capability of PROTECT to operate without distance and residual energy information, PROTECT can be applied to the network with sensor devices that have limited functionalities for the purpose of overcoming the energy hole problem.

CHAPTER 4

AN ENERGY-AWARE RANGE-EXTENSION CT PROTOCOL

4.1 Overview

In the previous chapter, we have seen the nontrivial performances of PROTECT in multi-hop WSNs. However, PROTECT has the following limitations: (i) PROTECT allows only the nodes in the same level to cooperate, (ii) PROTECT does not allow transmit-power control, and (iii) in order for PROTECT to work perfectly, conforming to the deployment method is required. Therefore, PROTECT can be highly effective in limited situations only, which motivates us to design a CT method that can be used without any constraints.

The basic concept of the new CT protocol is exactly the same as that of PROTECT. That is, a VMISO link is formed between cooperating nodes and a sink node using range-extension CT. However, unlike PROTECT, the residual-energy information of nodes, which is an important parameter of energy-aware routing protocols [32], [35], [34], [36], is used. The new distributed protocol developed in this section is referred to as "residual-energy-activated cooperative transmission (REACT)." In REACT, a node decides to do CT and selects its cooperators based on the residual energy instead of the probability of doing CT, and the analytical model developed in Section 3.3 is not necessary. Therefore, the deployment method of PROTECT is no longer required for REACT. Also, in REACT, a node can select its cooperators from its one-hop neighbors, and those neighbors can be in different levels. Moreover, the case where a node is able to control its transmit power can be handled in REACT.

The remainder of this chapter is organized as follows. We introduce assumptions and definitions required for developing the REACT protocol in Section 4.2. The REACT protocol is developed in Section 4.3, and the simulation results of REACT and non-CT protocols are presented in Section 4.4. Finally, we summarize this chapter in Section 4.5.

4.2 Assumptions and Definitions

In addition to the common assumptions and definitions introduced in Section 2.4, we make the following assumptions and definitions. A node can adjust its transmit power, and all nodes have the same maximum transmission range, $d_{\text{tx}}^{\max}$. Cooperative routing relies on a non-CT primary route, and the nodes in the primary route decide whether to do CT or not. An initiator should first do CT sharing before doing CT, and the VMISO communication is done after the CT sharing. A VMISO link can be formed between cooperating nodes and one of sink nodes.

Let us denote the distance between a node n_i and a sink node by $d_s(n_i)$ and the residual energy of node n_i by $E_{\text{re}}(n_i)$. The set of cooperating nodes is denoted by S_{CT}. Note that N_c (in Section 2.4.1) is equal to $|S_{\text{CT}}|$. If we denote the i-th cooperating node by $n_{\text{C},i}$ ($1 \leq i \leq N_c$), then $n_{\text{C},i} \in S_{\text{CT}}$. A transmitting node in a SISO communication link is denoted by n_{TX}, and the receiving node is denoted by n_{RX}.

The minimum required distance d_{req}, which is related to the minimum required transmit power of each cooperating node, is derived and explained in Appendix B. Using the fact that the intended VMISO receiver is the sink node, d_{req} can be rewritten using the variables defined in this section as

$$d_{\text{req}} = \left(10^{G(N_c)/10} \cdot \sum_{i=1}^{N_c} d_s(n_{\text{C},i})^{-\alpha} \right)^{-1/\alpha}, \tag{25}$$

where α is the path loss exponent. Note that d_{req} cannot exceed the maximum transmission range, $d_{\text{tx}}^{\max}$, because of the limitation of the transmission power (or range). Therefore, if $d_{\text{req}} > d_{\text{tx}}^{\max}$, CT cannot be used to directly communicate with the sink node. In other words, if the following condition does not hold,

$$d_{\text{req}} \leq d_{\text{tx}}^{\max}, \tag{26}$$

it is impossible to form a VMISO link between the cooperating nodes in S_{CT} and the sink node. Therefore, when a node cannot find S_{CT} that satisfies (26), the node cannot do CT, and the node should do non-CT.

Since REACT is an energy-aware routing scheme and many existing energy-aware routing schemes define the network lifetime as the time of the first node's death [35], [31], [32], this lifetime definition is used when evaluating REACT. Also, following [35] and [36], the lifetime performance is measured in terms of the number of data packets that successfully reach the sink node.

4.3 Developing the REACT Protocol

In this section, the REACT protocol, which uses the residual-energy information of nodes and CT's extended range to improve the network lifetime, is developed.

4.3.1 The Trigger for Using CT

Suppose n_{TX} and n_{RX} are two communicating nodes in a primary route. n_{RX} should transmit not only its own data but also the data received from n_{TX}. However, this does not necessarily imply that n_{RX} has more relay burden than n_{TX} because n_{TX} may have a lot of alternative paths to the sink node so that it does not have to use n_{RX} often, while a lot of data should go through n_{TX} (the nodes $n_{\mathrm{TX},1}$ and $n_{\mathrm{RX},1}$ in Figure 17). In any case, we can expect that a node with more relay burden should have less residual energy than the other as the time goes by. Therefore, in REACT, when n_{TX} has more residual energy than n_{RX}, n_{TX} tries to use range-extension CT to directly communicate with the sink node so that n_{RX} does not need to be used. To summarize, n_{TX} behaves as follows:

$$\text{if } E_{\mathrm{re}}(n_{\mathrm{TX}}) > E_{\mathrm{re}}(n_{\mathrm{RX}}), \text{ try to do CT, and}$$
$$\text{if } E_{\mathrm{re}}(n_{\mathrm{TX}}) \leq E_{\mathrm{re}}(n_{\mathrm{RX}}), \text{ do non-CT.}$$

Note that, with the above condition, the node (n_{TX}) that has a sink node as its next-hop node (n_{RX}) never triggers CT.

4.3.2 Selecting Cooperators for CT

When $E_{\mathrm{re}}(n_{\mathrm{TX}}) > E_{\mathrm{re}}(n_{\mathrm{RX}})$, n_{TX} tries to do CT and is in charge of selecting its cooperators. n_{TX} has to choose the desired number of cooperators, N_c, and form S_{CT} in a way to protect

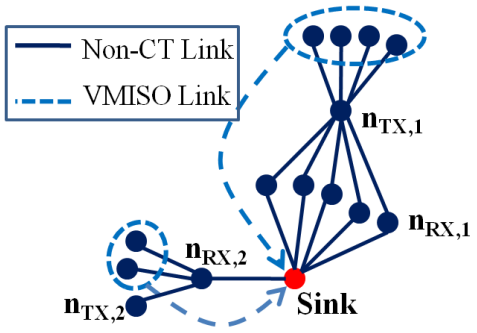


Figure 17: The illustration of using range-extension CT in the random deployment. In the random deployment, it is possible that the node that is not close to the sink node (such as $n_{TX,1}$) is highly burdened.

the highly burdened node and extend the network lifetime. In order to maximize the minimum residual energy left after the data transmission, it is obvious that n_{TX} must choose (i) the nodes with high $E_{re}(n_i)$'s and (ii) the nodes that give a low d_{req} (to reduce the transmit power). However, if we stick to a somewhat idealized cooperator-selection method, we may not be able to find S_{CT} that satisfies (26), and, when this happens, n_{TX} should do non-CT, which can nullify our overall purpose of protecting n_{RX} using CT. Note that the effectiveness of our approach comes from the fact that the energy consumption of the highly burdened node can be reduced by using the energies of other nodes which are going to be unused anyway if the highly burdened node dies early and the network is dead, rather than through optimizing the energy consumption. Therefore, a reasonable way to select the cooperators is considered for REACT, rather than the idealized cooperator selection.

The cooperator-selection process mainly consists of two parts: (i) selecting potential cooperators reasonably (the set of potential cooperators is denoted by S_p) and (ii) selecting desirable cooperators among the nodes in S_p and checking the possibility of communicating directly to the sink node using CT. In the first part of the cooperator-selection process, n_{TX} has to consider the nodes with high $E_{re}(n_i)$'s and the nodes that lead to a low d_{req}. As can be seen from (25), a low d_{req} can be achieved when the nodes with small $d_s(n_{C,i})$'s are selected. Therefore, it is evident that picking the nodes with high residual energy and short distance to the sink node is a good choice. Since "high" and "short" are subjective

matters, the average values, considering itself (n_{TX}) and all its neighbors, are calculated and used as a guideline. That is, n_{TX} first calculates the average residual energy, $E_{\mathrm{re}}^{\mathrm{avg}}$, and the average distance to the sink node, d_s^{avg}. Then, the nodes, n_i's, that satisfy the following two conditions are identified as potential cooperators:

$$E_{\mathrm{re}}(n_i) > \max(E_{\mathrm{re}}^{\mathrm{avg}}, E_{\mathrm{re}}(n_{\mathrm{RX}})), \tag{27}$$

and

$$d_s(n_i) < d_s^{\mathrm{avg}}, \tag{28}$$

where max(A,B) returns the largest of A and B. The reason why $E_{\mathrm{re}}(n_{\mathrm{RX}})$ is included in (27) is that there is no point in using the node having less residual energy than n_{RX} to protect n_{RX}. Note that considering both conditions (27) and (28) is close to the idealized selection method discussed above. After this process, S_{p} is formed (S_{p} is a subset of n_{TX}'s neighbors and does not include n_{TX}). Now, the second part of the selection process determines whether the direct communication to the sink node using CT is possible with the selected nodes in S_{p}. Note that, as mentioned earlier, the direct communication between cooperating nodes and the sink node is possible if and only if the condition in (26) holds.

Before we explain the second part, let us consider the case where the direct communication to the sink node is impossible (i.e., the condition (26) does not hold). In this case, it may still be possible to protect n_{RX} by increasing the number of potential cooperators, which can be done by loosening the conditions (27) and (28). Since the residual energy of a node is directly related to the lifetime of the node, (28) is removed, and the potential cooperators are selected using (27) only. A new S_{p} is formed, and the second part of the selection process (explained below) selects the nodes and checks the condition (26). If the condition (26) does not hold after the second selection process with the new S_{p}, n_{TX} does not do CT. The importance of this loosening of the conditions is discussed in Section 4.4.

We now discuss the second part of the selection process. The main objective of the second part is to choose the desirable cooperators among the nodes in S_{p}. Note that we

do not want N_c ($2 \leq N_c \leq N_d$) to be unnecessarily large, because this can increase the total energy consumption of doing CT. In order not to overuse cooperators, n_{TX} first sets $N_c = 2$, and do the following procedures until N_c reaches $\min(|S_{\mathrm{p}}|, N_d)$, where min(A,B) returns the smallest of A and B. n_{TX} selects $N_c - 1$ nodes from S_{p} that have the highest residual energies. The selected $N_c - 1$ nodes and n_{TX} form S_{CT}. For the nodes in S_{CT}, d_{req} is calculated using (25), and the condition (26) is checked. If the condition (26) holds, n_{TX} decides to do CT, and N_c nodes in S_{CT} perform CT. If the condition (26) does not hold, n_{TX} sets $N_c = N_c + 1$ and repeats the procedure. The resulting selection process, which combines the first and second parts, is summarized in Section 4.3.3.

Note that if n_{TX} already has the required information of the neighbors such as $d_s(n_i)$ and $E_{\mathrm{re}}(n_i)$, the procedures that we have introduced in this section and Section 4.3.1 can be done without any additional transaction with n_{TX}'s neighbors. This required information can be obtained by using a periodic message (HELLO message), which is widely used in non-CT and CT routing protocols [47], [35], [56]. Also, gathering the cooperators and doing CT can be done by using the cross-layer framework in [47]. However, REACT can be built on top of any cross-layer framework designed to use range-extension CT.

4.3.3 Summary of the REACT Protocol

When a source node needs to transmit its data to a sink node, it first establishes a primary route (or uses a pre-existing primary route). Then, along the primary route, when n_{TX} needs to transmit/relay a packet, it decides whether to do CT or not using the following procedure.

- **Step 0.** Set the variable *num_trial* to one. Also, $COND_A$ = (27) & (28), $COND_B$ = (27).

- **Step 1.** Check the residual energy, and if $E_{\mathrm{re}}(n_{\mathrm{TX}}) > E_{\mathrm{re}}(n_{\mathrm{RX}})$, calculate $E_{\mathrm{re}}^{\mathrm{avg}}$ and d_s^{avg}. Otherwise, decide to do non-CT and exit this procedure.

- **Step 2.** If *num_trial* is one, set *cond* to $COND_A$. If *num_trial* is two, set *cond* to $COND_B$. If *num_trial* is none of the above, decide to do non-CT and exit this

procedure.

- **Step 3.** Select potential cooperators (among its neighbors) satisfying *cond* and form S_{p}. Set $N_c = 2$ and proceed to the next step.

- **Step 4.** If $|S_{\mathrm{p}}| < N_c - 1$ or $N_c > min(|S_{\mathrm{p}}|, N_d)$, set $num\_trial = num\_trial + 1$ and go to Step 2. Otherwise, select $N_c - 1$ nodes from S_{p} that have the highest residual energies and save the selected nodes to the set S_{sel}. Set $S_{\mathrm{CT}} = \{n_{\mathrm{TX}}, S_{\mathrm{sel}}\}$ and proceed to the next step.

- **Step 5.** For the nodes in S_{CT}, calculate d_{req} using (25). Check the condition in (26). If (26) holds, decide to do CT and exit this procedure. If (26) does not hold, set $N_c = N_c + 1$ and go to Step 4.

Note that the calculation of d_{req} needs to be done at least once, and, in the worse case, $2 \cdot (N_d - 1)$ calculations are required. When n_{TX} decides to do CT, it becomes a CT initiator, and it is in charge of sharing its data with chosen cooperators and doing CT. When n_{TX} decides to do non-CT, n_{TX} forwards its data to the next-hop node (n_{RX}).

4.4 Simulation Results

In this section, the simulation results of REACT are provided. $d_{\mathrm{tx}}^{\mathrm{max}} = 20$m, and we use the diversity gains and path loss exponent defined in Section 2.4.2. The orthogonal diversity channel is assumed to be obtained by space-time block code (STBC) [9], and the maximum number of orthogonal channels is assumed to be four ($N_d = 4$). The energy model in [19][1] is used assuming 128 bytes of data. For diversity order larger than two, full rate STBC does not exist, and the best achievable rate for the diversity order of 3 and 4 is 3/4. This increases the energy consumption of doing CT for N_c=3 and 4, and the increased energy consumption is taken into account in the simulation. We consider the square-shaped

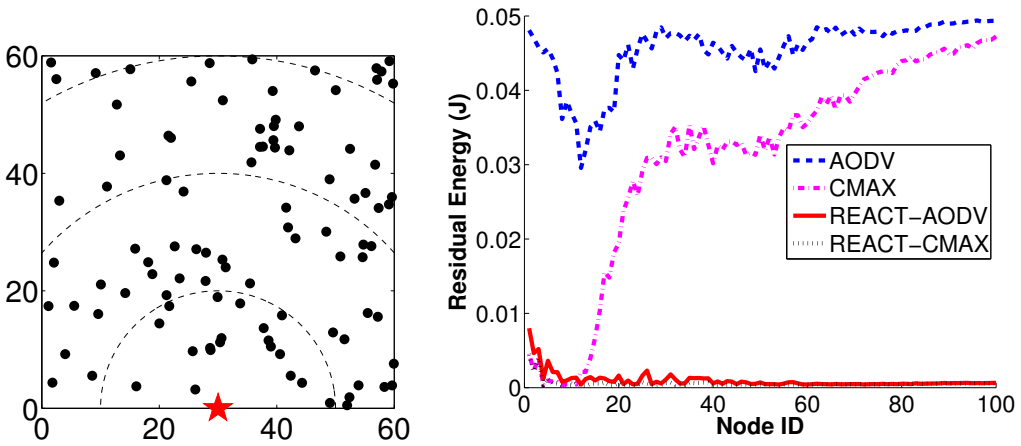
[1]The energy model in [19] has the transmit circuit energy of 45nJ/bit and the receive circuit energy of 135nJ/bit. For the radiated energy, we use $\beta \cdot d^{\alpha}$, where α is the path loss exponent and β=10^{-3}pJ/bit/m$^{\alpha}$.

networks with a single sink node located at the bottom center of the network. Nodes are randomly deployed, and each node has the initial energy of 0.05J. Any node except for the sink node can be a source at random.

For non-CT routing, two routing schemes are considered, which are (i) a shortest-hop routing, ad hoc on-demand distance vector (AODV) routing in [56] and (ii) CMAX introduced in Section 2.1.3. For REACT, we denote REACT using AODV (as the primary routing) by 'REACT-AODV,' and REACT using CMAX by 'REACT-CMAX.' The node ID of a node is assigned according to the proximity of the node to the sink node (the node closer to the sink node gets a lower node ID). The simulation is done until the first node dies. For each of the non-CT and REACT protocols, the simulation is repeated 20 times, and, in each simulation trial, nodes are randomly relocated except for the sink node.

We first provide the result for 60m×60m networks with 100 nodes. One network topology of 60m×60m networks is shown in Figure 18a, where the star indicates the location of the sink node and the dashed lines indicate C_i's defined in Section 3.2. The average residual energy of each node after the death of the first node is shown in Figure 18b. The figure shows that non-CT schemes (AODV and CMAX) cannot successfully use the energies of the nodes far away from the sink node (nodes with high node IDs). CMAX tries to evenly consume the energies of the nodes close to the sink node, but the residual energies of the nodes far away from the sink node are close to their initial energies, and these residual energies can be considered as wasted. Figure 18b clearly shows that both REACT-AODV and REACT-CMAX balance the energy consumption.

Balancing the energy consumption is meaningless if it does not lead to the lifetime extension. The average network lifetimes (the total number of packets received by the sink node) of non-CT schemes and REACT for 40m×40m, 50m×50m, 60m×60m, and 70m×70m networks are summarized in Figure 19. For all four network sizes, 100 nodes are used. As can be seen from Figure 19, the lifetime extension achieved through REACT is significant. For 60m×60m network, the extended lifetime when REACT-AODV is used



(a) 60m×60m network.

(b) The average residual energy of each nodes after the first node is dead.

Figure 18: The network topology and average residual-energy results for 60m×60m network.

is more than 11 times of the lifetime of AODV and is about three times of the lifetime of CMAX. Figure 19 also shows that the lifetime-extension factor increases as the network size grows. The reason for this is that, as the network size grows, the nodes directly connected to the sink node become a small fraction of the total, which leads to more severe energy-hole problem. Also, it can be observed that REACT-AODV can sometimes be better than REACT-CMAX. Since REACT already tries to balance the energy consumption, avoiding the nodes with low residual energies using CMAX does not always pay off because the route chosen by CMAX may use more nodes than the shortest-hop case (AODV), which increases the overall energy consumption of the route. Therefore, unlike the non-CT case where the energy-aware routing is always beneficial, in the case of REACT, there are cases where the energy-aware routing is unnecessary, and, because of this reason, neither CMAX nor AODV can be considered the most desirable choice for the primary routing of REACT.

The effectiveness of any CT scheme depends on the availability of cooperators. Therefore, REACT is affected by the node degree. When the node degree is small, some nodes may not gather enough cooperators to do CT, and this affects REACT's ability to avoid

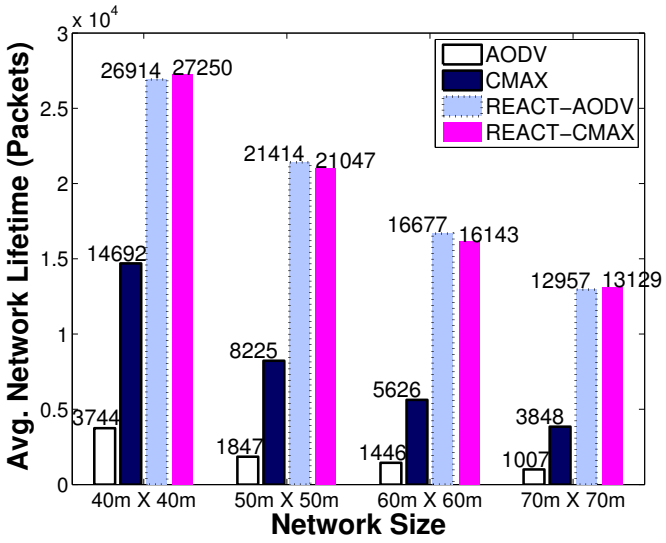
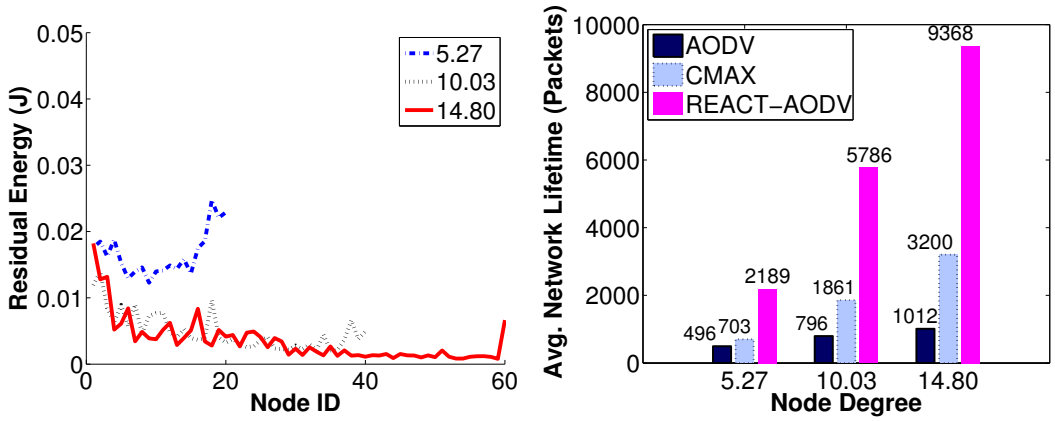


Figure 19: The average network lifetime for four different network sizes.

highly burdened nodes. To see this, we perform network simulations using different total numbers of nodes, 20, 40, and 60 for the network size of 60m×60m. The average node degrees for these three cases are 5.27, 10.03, and 14.80, respectively, and the results are shown in Figure 20. Figure 20a shows the average residual energy for REACT-AODV when the network is dead, which clearly shows that the node degree affects the REACT's ability of evenly (and completely) consuming the energy of nodes. Note that a low node degree also highly affects the performance of CMAX by limiting the path that can detour around the nodes with low residual energies, and therefore, as can be seen from Figure 20b, the lifetime of CMAX is relatively close to that of AODV and is very low when the node degree is 5.27. Since REACT still can partially balance the energy consumption when the node degree is 5.27, the lifetime-extension factor of REACT-AODV is more than three when compared with CMAX. The results of REACT-CMAX are similar to those of REACT-AODV, and the results are omitted.

In REACT, as summarized in Section 4.3.3, a node successively checks for two conditions when deciding potential cooperators (the set S_{p}): $COND_A$ and $COND_B$. To see the effects of these conditions, REACT is simulated using only one of the two conditions for 60m×60m network with 80 nodes. AODV is used for both non-CT and REACT. The

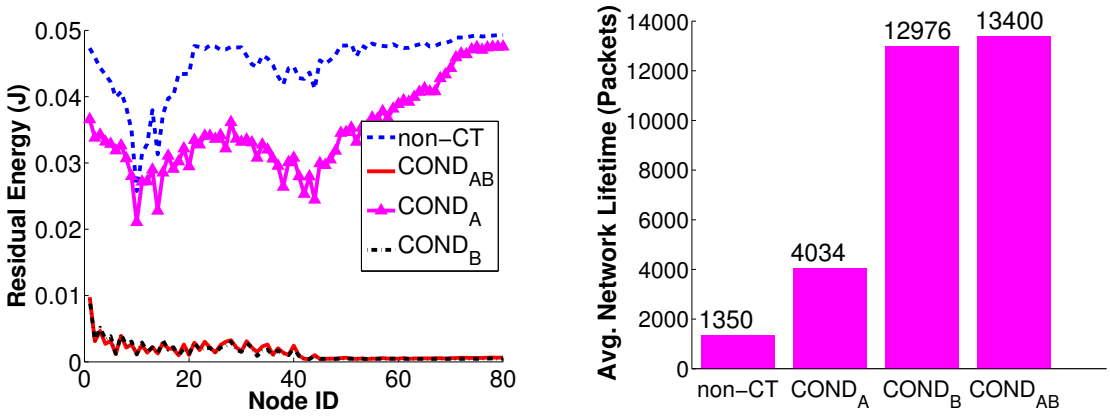


(a) The average residual energy of REACT-AODV.

(b) The average network lifetime.

Figure 20: The effect of different node degrees.

average residual energy of each node after the first node is dead is shown in Figure 21a, which shows four cases: (i) non-CT, (ii) REACT with $COND_A$ only, (iii) REACT with $COND_B$ only, and (iv) REACT with $COND_A$ and $COND_B$, denoted by $COND_{AB}$, which is the case presented in Section 4.3.3. Note that, as we have mentioned in Section 4.3.2, the condition $COND_A$ is close to the idealized selection. However, as can be seen from Figure 21a, using the condition $COND_A$ alone does a poor job in consuming the energy entirely. This is because this condition is too strict to gather enough cooperators to jump over the highly burdened node and communicate with the sink node directly. Failure to protect highly burdened nodes leads to a poor lifetime performance as shown in Figure 21b. The condition $COND_B$ alone uses most of residual energies (Figure 21a), which gives a larger lifetime value than $COND_A$ alone (Figure 21b). This shows that the effectiveness of our CT method stems mostly from the ability to avoid highly burdened nodes using the energies of the others, which are wasted if the highly burdened nodes die early. Still, the lifetime extension can be improved further when both conditions are used ($COND_{AB}$) as shown in Figure 21b.



(a) The average residual energy.

(b) The average network lifetime.

Figure 21: The effect of different conditions for CT decision.

4.5 Summary

In this chapter, we have designed a cooperative routing protocol, REACT, which regulates CT instances and selects cooperators, based on the residual energy. REACT can be applied to any existing non-CT routing protocol to significantly improve the network lifetime, and we have shown that the lifetime of the non-CT network that uses CMAX can be extended by a factor of more than three for large networks; this is because of the nature of the energy-hole problem and the REACT's ability to utilize the energy that cannot be efficiently used through non-CT routing. Moreover, REACT requires only the information of the one-hop neighbors to utilize a VMISO link, which makes the routing protocol simple and feasible.

CHAPTER 5

COOPERATIVE ROUTING IN ENERGY-HARVESTING NETWORKS

5.1 Overview

In previous chapters, we discussed using cooperative routing to extend the lifetime of the network that has no energy-harvesting capability. When nodes can harvest energy from energy sources, extending the lifetime of a network is less important than providing required services seamlessly. A node in energy-harvesting WSNs (EH-WSNs) also has a limited amount of energy that can be harvested, and therefore, once the node consumes all of its harvested energy, it cannot operate until its energy is replenished. Therefore, for multi-hop EH-WSNs, if the service requirement is too high, the energy-hole problem can still occur from time to time, which can cause frequent service failures. This motivates us to investigate the benefits of using range-extension CT for multi-hop EH-WSNs, which is discussed in this chapter.

We adopt a systematic approach to show the benefits of using range-extension CT for EH-WSNs. For a simple two-hop network, we show through analysis that the range-extension CT can provide better services compared to non-CT. Then, we introduce a method of determining the supportable service of EH-WSNs that use the optimal non-CT routing. By comparing REACT with a non-CT energy-aware routing protocol designed for EH-WSNs in [36] and the optimal non-CT case, the advantages of using range-extension CT for EH-WSNs are shown in this chapter.

The remainder of this chapter is organized as follows. Section 5.2 makes assumptions and definitions that are used in this chapter. In Section 5.3, we discuss the possibility of using range-extension CT to provide better services than non-CT for a simple two-hop network. A method of determining the supportable service of EH-WSNs that use the optimal non-CT routing is developed in Section 5.4. We evaluate our method and provide network

simulation results in Section 5.5, and we summarize this chapter in Section 5.6.

5.2 Assumptions and Definitions

We consider multi-hop EH-WSNs that run periodic-reporting applications. A node except for sink nodes has energy-harvesting capability and has to report sensed data periodically, with period T_{RP}. All nodes use the same transmit power for transmission. In EH-WSNs, the network operator wishes to run the network for several years or decades, and the amount of the desirable network lifetime will be referred to as the required service period, and we denote the required service period by T_{serv}. E_{Data}^{TX}, E_{Data}^{RX}, E_{Ack}^{TX}, E_{Ack}^{RX}, and E_{Sen} denote the energy consumptions for data transmission, data reception, acknowledgment transmission, acknowledgment reception, and sensing, respectively. We use P and T for the power and time values respectively that correspond to the energy value E. For example, the power and time values for E_{Data}^{TX} are denoted by P_{Data}^{TX} and T_{Data}^{TX}, respectively.

For CT networks, an initiator first sends its data to selected cooperators (CT sharing), and then, it performs CT with the selected cooperators to send the data to the VMISO receiver. A VMISO link can be formed between cooperating nodes and one of sink nodes. Since we are particularly interested in comparing the network protocols and there are many different ways to implement MAC layer, we do not consider any energy consumption regarding MAC except for the ACK packet communication for both non-CT and CT.

The "service requirement" or "required service" indicates that all data generated by source nodes should be delivered to sink nodes during T_{serv}. If the network can meet the service requirement for given network topology and T_{RP}, the service is *supportable*. The supportable service depends mainly on the network topology and T_{RP}, and the "upper bound" of the supportable service of a network is either the maximum size of the network for a given T_{RP} or the lowest value of T_{RP} for a given network topology that the network can support. A network is considered to provide a *better* service than the other if, given a fixed number of sink nodes, it can support either a larger network (provide better sensing

coverage) or smaller T_{RP} (more frequent data gathering).

5.3 Analysis of a Simple Two-hop Network

In this section, we discuss the possibility of using range-extension CT to provide better services than non-CT through a simple analysis. Let us denote the harvested energy during T_{RP} by E_{EH}. In this section, we assume that E_{EH} is constant for all nodes. During T_{RP}, a node must consume energy for sensing and transmitting (its own sensed data), denoted by E_{ST}, and $E_{ST} = E^{TX}_{Data} + E^{RX}_{Ack} + E_{Sen}$. Let E_A=$E_{EH} - E_{ST}$ be the available energy that each node can use during T_{RP} (because E_{ST} must be used during T_{RP}). We assume that nodes do not have any energy at the beginning, and they have to rely on the harvested energy to operate. When a node is i hops away from the sink node, we say that the node is in "Level i" (as we have used in Chapter 3). Here, the hop count is determined based on the SISO communication range.

We consider a simple two-hop network in Figure 22, where there are n_1 nodes in Level 1 and n_2 nodes in Level 2. In Figure 22, we assume that each node in Level 2 can communicate with any node in Level 1 and Level 2. We also assume that the incoming traffic (relay traffic) of the nodes in Level 1 is distributed as equally as possible.

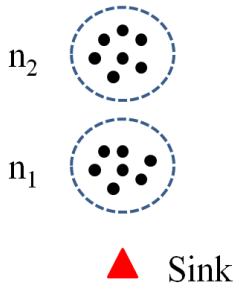


Figure 22: A simple two-hop network. There are n_1 nodes in Level 1 and n_2 nodes in Level 2. Any nodes in Level 1 and Level 2 can talk to each other.

In the case of non-CT, some nodes in Level 1 must relay $\lceil \frac{n_2}{n_1} \rceil$ messages, whereas the rest can relay only $\lfloor \frac{n_2}{n_1} \rfloor$ messages (because the incoming traffic is almost equally distributed). Therefore, in the worst case, a node in Level 1 must relay $\lceil \frac{n_2}{n_1} \rceil$ messages. Now, let us

consider the number of relay messages that a node in Level 1 can handle during the time interval $[0, T_{RP}]$. We note that relaying is a distinct operation from "sensing and transmitting." When a node relays, it consumes $E_{Relay} = P^{TX}_{Data}T^{TX}_{Data} + P^{RX}_{Data}T^{RX}_{Data} + P^{TX}_{Ack}T^{TX}_{Ack} + P^{RX}_{Ack}T^{RX}_{Ack}$. Therefore, a node can relay $\lfloor E_A/E_{Relay} \rfloor$ messages, and this number should be larger than $\lceil \frac{n_2}{n_1} \rceil$ (the worst case), which leads to the following condition for non-CT:

$$\left\lceil \frac{n_2}{n_1} \right\rceil \leq \left\lfloor \frac{E_A}{E_{Relay}} \right\rfloor \triangleq K. \tag{29}$$

Now, let us consider the case where the nodes in Level 2 form a VMISO link directly to the sink node by using range-extension CT so that the relay burden of the nodes in Level 1 can be reduced. We assume that two cooperating nodes are enough to form a VMISO link to the sink node. If we assume that n_x nodes ($n_x \leq n_2$) in Level 2 initiate CT for their sensed data so that n_x packets do not have to be relayed by the nodes in Level 1, then the nodes in Level 1 have to take care of only $\lceil \frac{n_2 - n_x}{n_1} \rceil$ relay messages during T_{RP} in the worst case, which leads to the following condition for CT:

$$\left\lceil \frac{n_2 - n_x}{n_1} \right\rceil \leq K. \tag{30}$$

Let us assume that $n_x \leq \frac{1}{2}n_2$ so that the sets of initiators and cooperators can be two mutually exclusive sets. In other words, there will be n_x initiators and n_x cooperators among n_2 nodes in Level 2, and none of them are the same node. An initiator that decides to do CT for transmitting its own sensed data, has to consume its energy for (i) data sharing, denoted by E^{TX}_{SH}, which is $P^{TX}_{Data}T^{TX}_{Data} + P^{RX}_{Ack}T^{RX}_{Ack}$, and (ii) doing CT with its cooperator, denoted by E^{TX}_{CT}, which is $P^{TX}_{Data}T^{TX}_{Data} + P^{RX}_{Ack}T^{RX}_{Ack}$. Note that, since the initiator sends its sensed data through CT, E^{TX}_{CT} should be considered as already included in E_{ST} of E_A, and therefore, an initiator has to use an E^{TX}_{SH} out of its available energy. The cooperator that sends its sensed data using non-CT and has to do CT for the initiator's sensed data additionally must consume its energy for (i) receiving CT sharing message, denoted by E^{RX}_{SH}, which requires $P^{RX}_{Data}T^{RX}_{Data} + P^{TX}_{Ack}T^{TX}_{Ack}$ and (ii) doing CT with the initiator, which requires E^{TX}_{CT}. Therefore, a cooperator has to use $E^{RX}_{SH} + E^{TX}_{CT}$ out of its available energy, and this is the worst case energy

consumption for the nodes in Level 2. Note that each cooperator handles exactly one CT (n_x CT instances are distributed to n_x cooperators), so $E_{SH}^{RX} + E_{CT}^{TX} \leq E_A$ should hold, which can be rewritten as

$$1 \leq \frac{E_A}{E_{SH}^{RX} + E_{CT}^{TX}} = \frac{E_A}{E_{Relay}}. \tag{31}$$

As long as (30) and (31) hold, the nodes in Level 2 can successfully perform n_x CT instances where $n_x \leq \frac{1}{2}n_2$. Note that $\lceil \frac{n_2}{n_1} \rceil \geq 1$, which means that the condition for non-CT in (29) already implies (31). It is also obvious that if (29) holds, so does (30). Therefore, if (29) holds, which means that non-CT can support the required service, doing CT for n_x times in Level 2 is also possible.

Now, let us consider the number of nodes in Level 2 that the network can support for CT and non-CT. From (29), we can get $n_2 \leq Kn_1$ for non-CT, and, from (30), we get $n_2 \leq Kn_1 + n_x$ for CT. This means that, for given K and n_1, CT can support n_x more nodes compared to non-CT where $n_x \leq \frac{1}{2}n_2$. In other words, in CT networks, $100 \times n_2/\{2(n_1+n_2)\}$ percent more nodes can be supported compared to non-CT networks for the network in Figure 22.

Let us consider the reporting period T_{RP}. Low T_{RP} means that the sensed data can be gathered more frequently, which lets us have more up-to-date status. We plot K's for different T_{RP}'s (in seconds) in Figure 23 using the values in Table 7, which are measured values of real energy harvesters and sensor devices [57], [58]. More specifically, the measurements are obtained by using Tyndall motes with temperature/humidity sensors, Sensirion SHT71, and dual-axis accelerometers, Analog Devices ADXL250 [57]. Also, $P_{EH} = (P_{PV} \times \eta_{MPPT} - P_{Leak}) \times \eta_{SR}$, where $P_{PV} = 382\mu W$ is the maximum output of a Sanyo AM1815 solar panel under 500 Lux fluorescent light, $\eta_{MPPT} = 90.5\%$ is the efficiency of the maximum power-point tracker (MPPT), $P_{Leak} = 35\mu W$ is the average leakage power of Maxwell BCAP0005 supercapacitor, and $\eta_{SR} = 51\%$ is the average conversion efficiency of TI TPS61220 switching regulator [58]. As can be seen from Figure 23, K is a monotonically increasing function of T_{RP}, and, in order to support low T_{RP}, K should be low.

Table 7: Power and time values [57], [58].

Power (mWatt)		Time (msec)	
$P^{TX}_{Data} = 131.4$	$P^{TX}_{Ack} = 111.1$	$T^{TX}_{Data} = 35$	$T^{TX}_{Ack} = 0.45$
$P^{RX}_{Data} = 174.3$	$P^{RX}_{Ack} = 133.3$	$T^{RX}_{Data} = 35$	$T^{TX}_{Ack} = 0.45$
$P_{Sen} = 53.3$	$P_{EH} = 0.1585$	$T_{Sen} = 7.5$	

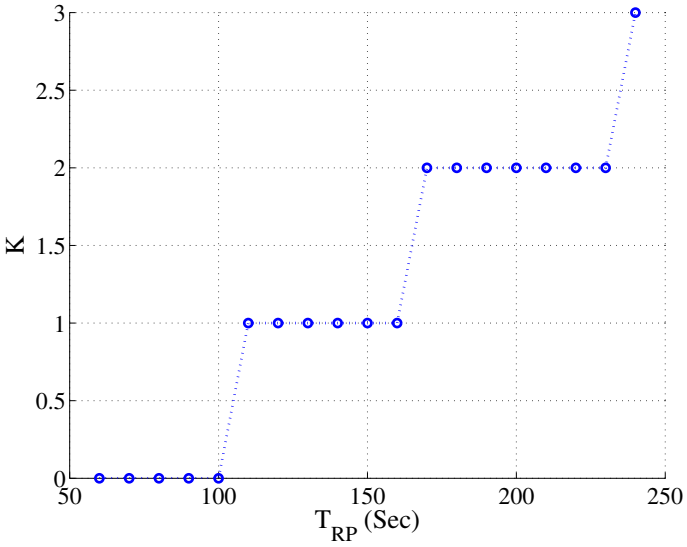


Figure 23: The values of K for different reporting periods.

For, K=0, (29) and (31) are violated, and therefore, both CT and non-CT networks cannot support $T_{RP} \leq 100$. For K=1 ($110 \leq T_{RP} \leq 160$), non-CT should satisfy $\lceil \frac{n_2}{n_1} \rceil \leq 1$ and CT should satisfy $\lceil \frac{n_2 - n_x}{n_1} \rceil \leq 1$ (when K=1, (31) readily holds). Therefore, when $n_2 \leq n_1$, or equivalently, $\lceil \frac{n_2}{n_1} \rceil = 1$, both CT and non-CT networks can support T_{RP}=110 secs. When $\frac{1}{2}n_2 \leq n_1 < n_2$, or equivalently, $\lceil \frac{n_2}{n_1} \rceil = 2$, (29) forces K to be 2 and non-CT networks can only support as low as T_{RP}=170 secs, whereas, since $\lceil \frac{n_2 - n_x}{n_1} \rceil$ is still less than or equal to one because $n_x \leq \frac{1}{2}n_2$, CT can still have K=1 meaning CT network can still support T_{RP}=110 secs, which is shorter than 170 secs of non-CT networks. Therefore, in the case of CT networks, it is possible that data can be reported more frequently than non-CT networks.

5.4 Determining Supportable Services of Non-CT EH-WSNs

In this section, we introduce a method of determining the supportable service of EH-WSNs that use the optimal non-CT routing. The reason why this method is necessary is that it can determine the supportable service of the ideal non-CT networks, and the determined supportable services of non-CT networks can be compared with the supportable services of CT networks. If any cooperative routing can provide better services than the ideal non-CT networks, we can safely claim that CT networks can provide better services than non-CT networks. Note that one can select an existing non-CT routing protocol and run network simulations to determine the supportable services. However, relying on an existing routing protocol may not be sufficient to determine the true supportable services of non-CT networks because there is no guarantee that the selected routing protocol is performing optimally under given circumstances.

To develop the method, we use the idea of [31] introduced in Section 2.1.2, and we use all the variables defined in Section 2.1.2. Note that the "data unit" in Section 2.1.2 is the "data packet" in this section. Note that, in WSNs, D (the set of destination nodes) is the set of sink nodes, and, because sink nodes are not energy constrained, $D \subset E$. The total number of source nodes in the network will be denoted by N_s ($= |A| - |D|$). We define e_i^{gen} as the energy consumption for generating a data packet.

In the case of energy-harvesting networks, we are no longer looking at the lifetime maximization, and whether one can get the required services in the service period (T_{serv}) is important. We use the fact that an optimal routing for EH-WSNs should be able to deliver all required data successfully to the sink nodes during T_{serv}. The total number of received data during T_{serv} for non-CT networks can be expressed as

$$\sum_{k:k\in D} \sum_{j:k\in S_j} n_{jk}, \tag{32}$$

where, n_{ij} is redefined as the total number of data packets transmitted from Node i to Node j until T_{serv} instead of T ($T = T_{\text{serv}}$). For periodic-reporting applications, we want the

sink nodes to receive a total of $N_s \cdot \lfloor T_{\text{serv}}/T_{RP} \rfloor$ packets at least[1]. In order to eliminate the ambiguity of the required number of packets to be received, we assume that T_{serv} is a multiple of T_{RP} from now on, and, for the service time of T_{serv}, a total of $N_s \cdot T_{\text{serv}}/T_{RP}$ packets should be delivered to the sink nodes.

The energy-constraint condition of Node i when energy harvesting is possible can be expressed as follows:

$$(T_{\text{serv}} \cdot Q_i) \cdot e_i^{\text{gen}} + \sum_{j \in S_i} e_{ij}^{\text{TX}} \cdot n_{ij} + \sum_{j:i \in S_j} e_{ji}^{\text{RX}} \cdot n_{ji} - (T_{\text{serv}} \cdot H_i) \leq E_i^{\text{I}}, \quad \forall i \in A - E, \tag{33}$$

where H_i is the time-averaged harvesting rate of Node i in Joules/sec, and $T_{\text{serv}} \cdot H_i$ represents the overall harvested energy during T_{serv}. Note that maximizing (32), when combined with the energy-constraint condition and the flow-conservation condition, captures the optimal routing behavior for energy-harvesting networks.

From the above discussions, we formulate the LP problem for non-CT EH-WSNs as

$$\begin{aligned} &\text{Maximize} \sum_{k:k \in D} \sum_{j:k \in S_j} n_{jk} \\ &\text{s.t. (2), (33),} \\ &\sum_{j:i \in S_j} n_{ji} + T_{\text{serv}} \cdot Q_i = \sum_{j \in S_i} n_{ij}, \qquad \forall i \in A - D, \end{aligned} \tag{34}$$

and if (32) = $N_s \cdot T_{\text{serv}}/T_{RP}$, the optimal non-CT routing scheme can support the required service during T_{serv}. Note that (33) does not consider the capacity limit of the energy storage device, and we argue that not considering the capacity limit does not critically harm the objective of the above method because of the following reasons. Whether (32) can meet $N_s \cdot T_{\text{serv}}/T_{RP}$ or not is dependent on the nodes that get relatively low energy reserves (compared to other nodes) as time goes by. In other words, (32) cannot have $N_s \cdot T_{\text{serv}}/T_{RP}$ when the nodes with relatively low energy reserves do not have energy to complete all their sending tasks. Those nodes usually cannot reach their capacity limits,

[1]The total number of received packets should be between $N_s \cdot \lfloor T_{\text{serv}}/T_{RP} \rfloor$ and $N_s \cdot \lceil T_{\text{serv}}/T_{RP} \rceil$ depending on T_{serv} and the time each node reports its data.

and therefore, they should not be affected by having storage capacity limit or not. Also, even if there are situations that those nodes reach their capacity limits, the above method can still give the upper bound on the supportable services because it assumes the case of an ideal storage device. Note that, unlike T, T_{serv} is not an LP variable. Also, for periodic-reporting applications, $Q_i = 1/T_{RP}$ for all $i \in A - D$. This means that the above LP formulation already implies that each node is able to generate T_{serv}/T_{RP} ($= T_{\text{serv}} \cdot Q_i$) packets and transmit them to sink nodes during T_{serv}, and, when the LP solution exists, (32) $= N_s \cdot T_{\text{serv}}/T_{RP}$. Therefore, when the solution of the above LP exists, there exists an optimal non-CT routing scheme that enables the non-CT network to support the required service, and when the solution does not exist, non-CT networks cannot support the required service. Note that this also suggests that one can determine whether non-CT can support the required service or not by checking whether a feasible solution exists for the linear equalities and inequalities of (2), (33), and (34), which does not necessarily require LP formulation and solution.

We will refer to the method that can determine the supportable service of non-CT EH-WSNs using the above conditions as the "condition-based decision (CBD)" from now on. Also, the conclusion drawn from CBD will be referred to as the "CBD result." Note that the LP formulation in [31] (introduced in Section 2.1.2) can also provide the optimal lifetime for time-varying data generation rate because it considers total number of generated data during network lifetime ($T \cdot Q_i$) in its problem formulation (in (4)), and the formulation is not restricted to periodic-reporting applications. Likewise, our CBD introduced in this section can be used for any applications including periodic-reporting applications.

5.5 Evaluation

In this section, we verify our claim that range-extension CT can provide better services than non-CT in EH-WSNs by using CBD and performing network simulations.

5.5.1 Models and Parameters

For network simulations, we use the online non-CT routing protocol, E-WME[2], in [36] and the online CT routing protocol, REACT, introduced in Section 4.3, and we observe the throughput and energy behaviors of these two protocols. For the non-CT primary routing for REACT, we use E-WME.

We consider the network having one sink node. Note that neither CBD in Section 5.4, nor REACT, nor E-WME requires having a single sink in the network. We use the values in Table 7 for our network simulations assuming a fixed transmit power. Neither CBD, nor REACT nor E-WME is restricted to the fixed transmit power case, however, we use a fixed transmit power because, as we have mentioned earlier in Section 2.3, the radiated energy for existing sensor radios usually has a minor contribution to the total energy consumption.

All nodes have the same maximum SISO transmission range of 25m ($d_{\text{tx}}^{\max}$=25m), and a SISO link exists from Node i to Node j if Node j is within $d_{\text{tx}}^{\max}$ from Node i. Note that this makes links deterministic, and we can define n_{ij} for CBD. The orthogonal diversity channel for CT is obtained by STBC [9], and the maximum number of orthogonal channels is assumed to be four. When calculating d_{req} for REACT, we use the diversity gains and path loss exponent defined in Section 2.4.2.

For CBD, e_i^{gen}=E_{Sen}, e_{ij}^{TX}=E_{Data}^{TX} + E_{Ack}^{RX}, e_{ji}^{RX}=E_{Data}^{RX} + E_{Ack}^{TX} are used. As mentioned in Section 4.4, for diversity order larger than two, full rate STBC does not exist, and the best achievable rate for the diversity order of 3 and 4 is 3/4. Therefore, when the number of cooperating nodes are 3 or 4, we use $T_{Data}^{TX} \cdot 4/3$ (and corresponding E_{Data}^{TX}) for the network simulations of REACT.

In the network simulation, we assume that the storage capacity limit of a node is 5(J), and we make the sensing tasks to be well scheduled so that each sensed data is generated in different time incurring no collision or interferences for both CT and non-CT networks.

[2]We choose the E-WME algorithm because it has provably good performance (asymptotically optimal) and is specifically designed for energy-harvesting networks. E-WME uses the cost metric that includes storage capacity, harvesting rate, residual energy, and energy consumption, and the minimum-cost route is selected.

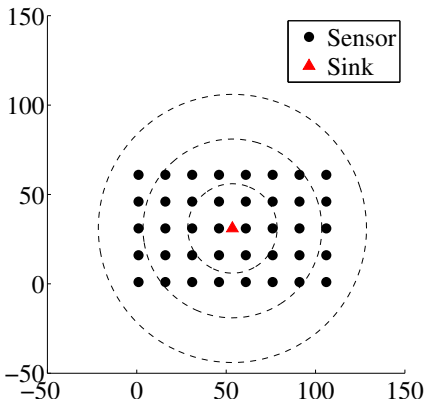
More specifically, for the k-th reporting interval $[(k-1)\cdot T_{RP}, k\cdot T_{RP}]$, Node i senses and reports its data at $(k-1)\cdot T_{RP}$ +$(i-1)\cdot T_{RP}/N_s$, however, when the node cannot sense or report its data at this time because it does not have enough energy, it completes its sensing task as soon as it has enough energy. At time T_{serv}, no node does the sensing task because the required service time is finished.

The initial energy of each node is assumed to be 0.02(J). We set the initial energy value to be low so that we can capture the impact of the energy harvesting earlier in the network simulations. We make the service demanding to see the upper bound of the supportable services. Because of the highly demanding service and nodes having small initial energy, we are able to detect the failure of meeting the service requirement earlier in the network simulations. Therefore, we set T_{serv}=24 hours and use it for CBD and network simulations. We consider both grid networks and randomly deployed networks. In the case of the grid networks, we consider a square-grid network with 15m minimum node spacing and a sink node at the center of the network. For randomly deployed networks, we consider square-shaped networks with a sink node located at the bottom center of the network. We assign the Node ID of a node according to the proximity of the node to the sink node (the node closer to the sink node gets a lower Node ID).

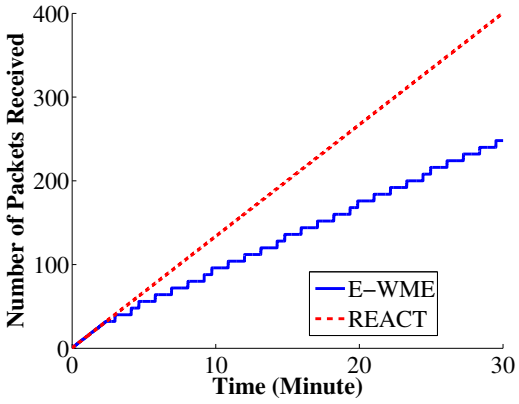
5.5.2 Results

We first look at the case of the constant harvesting rate. In this case, we assume that the harvesting rate is P_{EH} in Table 7, and the energy is constantly replenished. Let us first consider the 5×8 grid network shown in Figure 24a. The dashed circles in Figure 24a indicate Levels 1-3. The network simulation results of E-WME and REACT for this network are summarized in Figure 24.

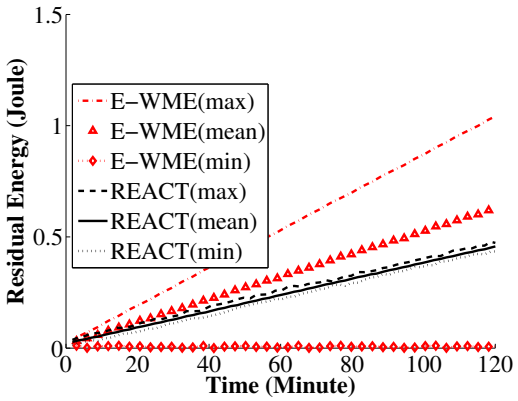
Figure 24b shows the number of packets received by the sink node over time when $T_{RP} = 3$ minutes. In this case, all the sensed data from 40 nodes should be delivered to the sink node every 3 minutes, and it can be seen from Figure 24b that REACT serves the network as desired, whereas E-WME cannot, which indicates that $T_{RP} = 3$ minutes



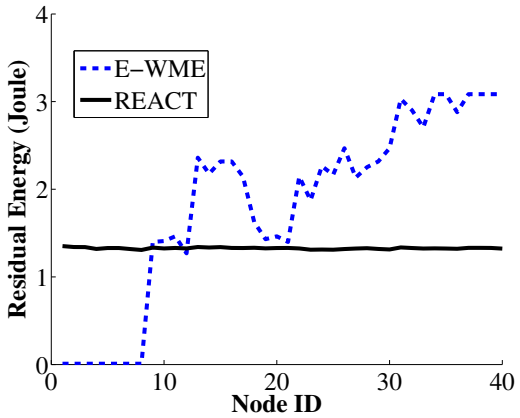
(a) Network Topology.



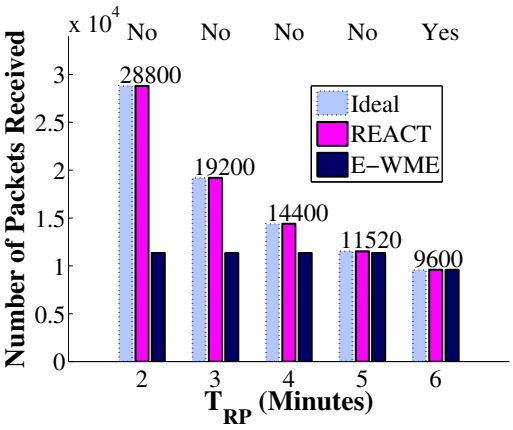
(b) Received packets vs. time. T_{RP} = 3 minutes.



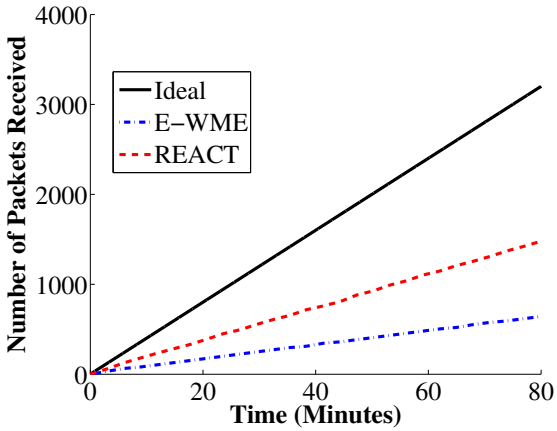
(c) Residual energy vs. time. T_{RP} = 3 minutes.



(d) Residual energy right after six hours. T_{RP} = 3 minutes.



(e) Received packets vs. reporting period after the simulation ends.



(f) Received packets vs. time. T_{RP} = 1 minute.

Figure 24: Network simulation results and topology of 5x8 grid network.

is too demanding for E-WME. Note that in the beginning of the simulation, E-WME and REACT have the identical performance because each node has its initial energy to spare and does not shut down in the beginning even though it overuses its energy. Figure 24c shows the mean, maximum, and minimum residual energies of nodes over time. As can be seen from the figure, the residual energy averaged over the network increases for both E-WME and REACT because energy is constantly replenished, however, it can be observed that the difference between the maximum and minimum energies of E-WME is very large indicating a significant energy imbalance in the network, whereas REACT's minimum energy value is relatively large and close to the maximum value, showing a balanced energy consumption. Figure 24d shows the residual energy snapshot of each node right after six hours. As can be seen from the figure, in the case of E-WME, the nodes close to the sink node (nodes with low Node IDs) have very low residual energies compared to the others. This explains the poor performance of E-WME in Figure 24b; the nodes close to the sink node run out of their energies frequently, and the sensed data cannot reach the sink node until those nodes replenish their energies. Figure 24e shows the number of packets received by the sink node after the simulation ends for five different reporting periods. Here, an imaginary "ideal" case where any sensed data can immediately reach the sink without any delay is considered also, and the number on top of each bar group shows the value of the ideal case. For all T_{RP}'s in the figure, REACT, representing the range-extension CT case, provides the same throughput results as the ideal case, whereas, in the case of E-WME, representing the non-CT case, deviates from the ideal when $T_{RP} \leq 5$ indicating that non-CT cannot meet the service requirement for $T_{RP} \leq 5$ (For $T_{RP} = 5$, E-WME has the value of 11352). The residual-energy trends over time for T_{RP}=2, 4, and 5, although they are omitted in Figure 24, are all similar to the case of $T_{RP} = 3$ in Figure 24c where the minimum energy of REACT goes up as the time goes by and the minimum energy of E-WME is nearly zero over time. Note that observing the minimum residual energy trend over time along with the throughput result can also be used to determine the supportable

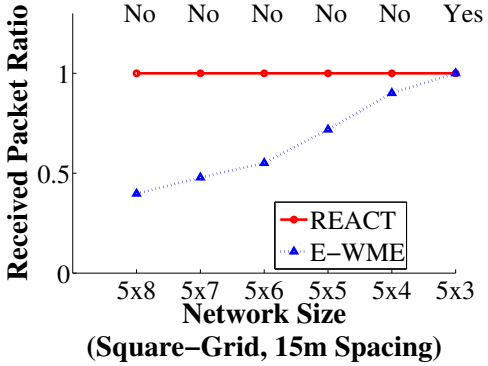
services. That is, when the minimum residual energy has an increasing trend over time and the throughput behavior matches with the ideal case, we can safely conclude that the service can be supported for an infinite amount of time as long as hardware does not fail. On the other hand, we can conclude that the service cannot be supported when the network fails to match the ideal throughput and has the minimum residual energy over time that decreases to and stays around zero (as in E-WME in Figure 24c). From now on, determining the supportable service using the minimum residual energy and throughput results will be referred to as the "simulation-based decision (SBD)." Using SBD, we can conclude that REACT can successfully provide the required service for $2 \leq T_{RP} \leq 5$, whereas E-WME cannot.

If we are completely sure that E-WME is operating optimally for given circumstances, we can conclude that non-CT networks can only support $T_{RP} = 6$ minutes by using SBD. To make it sure, we can use CBD, and, Figure 24e, in addition to simulation results, shows CBD results at the top of the graph; 'No' indicates that non-CT cannot provide the service successfully according to CBD, and 'Yes' means the opposite. CBD concludes that non-CT networks can only support $T_{RP} = 6$ minutes, and this result is consistent with that of the SBD for E-WME (representing non-CT), which shows the usefulness of CBD in determining the supportable service.

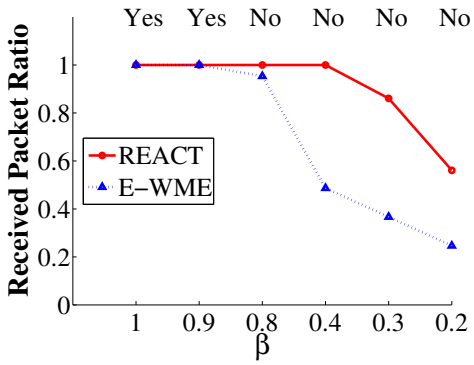
From the above discussion, CT networks can have less T_{RP}, meaning that more frequent data reporting is possible for CT networks compared to non-CT networks. The network simulation results showing the number of packets received by the sink node over time when $T_{RP} = 1$ minute is shown in Figure 24f. As can be seen from the figure, both REACT and E-WME underperform the ideal case because the service ($T_{RP} = 1$ minute) is too demanding. However, even in this case, the throughput of REACT is better than that of E-WME. The reason for this is that the non-CT network has to wait for the nodes that are close to the sink node to replenish their energies when the network is disconnected, whereas CT network may wait less than that because it has an option to use a VMISO link instead

of using the nodes close to the sink node.

Figure 25 shows the other benefits of using range-extension CT when T_{RP} is fixed. Figure 25a shows the network simulation results for different network sizes (e.g., network size of 5×7 means total 35 nodes in the square-grid network) when T_{RP} = 2 minutes. Here, the received packet ratio, measured after simulation ends, is the ratio of the number of received packets when a certain routing protocol is used to the number of received packets of the ideal case. SBD (we omit the results of the minimum residual energy over time) concludes that REACT can support all network sizes considered in Figure 25a and E-WME can only support 5×3 network, which is consistent with CBD results shown at the top of Figure 25a. Figure 25b shows the case where we have used the harvesting rate of $\beta \cdot P_{EH}$ instead of P_{EH} for 5×8 network with T_{RP} = 6 minutes, which is intended to observe the effect of poor energy-harvesting environments. The SBD concludes that REACT can support as low as β = 0.4, whereas E-WME can support as low as β = 0.9 only, showing that CT can operate the network successfully in a poor energy-harvesting environment (where the harvested energy is relatively weak) where non-CT cannot support. Again, SBD results for E-WME are consistent with CBD results.



(a) Received packet ratio vs. network size. T_{RP} = 2 minutes.



(b) Received packet ratio vs. β. 5×8 network. T_{RP} = 6 minutes.

Figure 25: The results of network simulations and CBD for grid networks.

The benefits of using CT are not restricted to a certain network topology. For example, we can also observe the effectiveness of using CT for randomly deployed networks. In the

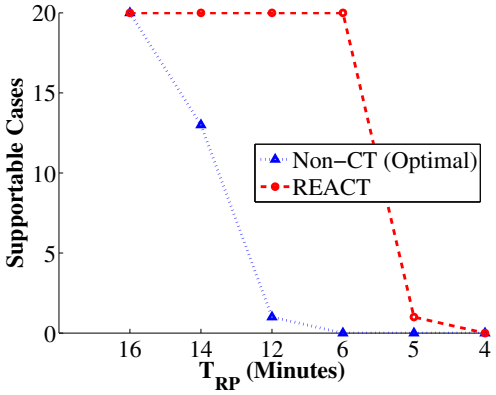
case of the random deployment, even when the same number of nodes and T_{RP} are used, the supportable service depends on how nodes are located. Therefore, we do CBD and SBD for 20 different random deployments given a network size and provide the results. For, 50m×50m with total 50 nodes (excluding the sink node) and $T_{RP} = 2$ minutes, the non-CT network is able to support 6 cases out of 20, and, when REACT is used, the network is able to support all 20 cases considered. When we change T_{RP} to 3 minutes, the non-CT network can support 17 cases out of 20, and the CT network that uses REACT can support all 20 cases considered. For, 75m×75m with total 60 nodes (excluding the sink node) and $T_{RP} = 3$ minutes, the non-CT network cannot support any cases out of 20, and, when REACT is used, the network can support all cases considered. This again shows that CT networks can provide better services compared to non-CT networks for randomly deployed networks. For all cases that we considered for randomly deployed networks, SBD results have shown to be consistent with CBD results.

Now, we consider the case where the harvesting rate of each node is time-varying and each node has different P_{EH}'s. Since the energy source for our harvesting model is the indoor light [58], we do the following. We use a Bernoulli random variable, γ_j, with success probability of 0.5 to represent if the light is on or not in the j-th minute. To model light intensity, for Node i, we generate an independent random variable, α_i, uniformly distributed over [0.5,1]. For the i-th node in the j-th minute, the harvesting rate is $\gamma_j \cdot \alpha_i \cdot P_{EH}$. This gives the average harvesting rate of $0.375 \cdot P_{EH}$, which is much less than P_{EH} of the constant harvesting rate case. The resulting overall harvested energy is also used for $T_{\text{serv}} \cdot H_i$ in (33)[3]. Square-grid networks are again used, and for fixed network size and T_{RP}, we do 20 trials (20 different harvesting conditions), and, for each trial, we determine whether CT and non-CT networks can successfully support the required service or not.

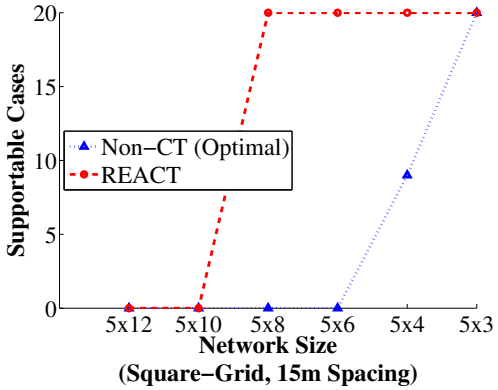
Figure 26 shows the number of cases (out of 20 cases) that the required service is

[3]Unlike the case of the constant harvesting rate, where the harvesting rate is predictable, the amount of the harvested energy in the time-varying case is usually unknown in advance. We can use a known $T_{\text{serv}} \cdot H_i$ value in (33) because CBD is for capturing the ideal case.

determined to be supportable. The cases of the optimal non-CT are determined using CBD, and the case of REACT is determined using SBD. The case of E-WME is omitted here because it does not outperform the optimal non-CT case. Figure 26a shows the number of supportable cases when different T_{RP}'s are used for 5×8 network. Here, when a routing scheme can support all 20 cases, we say that the scheme can "sufficiently" support the service. The optimal non-CT routing scheme can sufficiently support T_{RP} = 16 minutes only, whereas REACT can sufficiently support much lower T_{RP} (=6 minutes) showing CT's advantage over non-CT in data gathering. Note that the time-varying harvesting rate that we consider provides much less harvested energy than the case of the constant harvesting rate we have considered earlier, and we have seen that the non-CT network can support 5×8 with T_{RP} = 6 minutes when the constant harvesting rate is used. This again shows that, CT network can survive using less harvested energy compared to non-CT networks. Figure 26b shows the number of supportable cases for different network sizes when T_{RP} = 6 minutes, and the figure clearly shows that larger networks can be sufficiently supported for CT networks compared to non-CT networks; REACT can sufficiently support 5×8 network, which is larger than 5×3 network of the optimal non-CT case.



(a) Number of supportable cases vs. T_{RP}. 5×8 square-grid network.



(b) Number of supportable cases vs. network size. T_{RP} = 6 minutes.

Figure 26: The results of network simulations and CBD for time-varying harvesting rate. Grid networks.

5.6 Summary

In this chapter, we studied the advantages of using range-extension CT in multi-hop EH-WSNs. For a two-hop toy network, we have shown the possibility of using range-extension CT to support more nodes and more frequent data gathering compared to non-CT. Then, we have devised a method of determining the supportable service of EH-WSNs that use the optimal non-CT routing. Through evaluations of our decision method and network simulations, we showed the advantages of CT networks (that uses range-extension CT) over non-CT networks, which are (i) CT network can support larger networks, which provides better sensing coverage, (ii) CT network can gather data more frequently, which allows one to have more up-to-date status, and (iii) CT network can operate with lower harvested energy giving better chances of survival when nodes cannot harvest the expected amount of energy for some reason.

CHAPTER 6

LIFETIME OPTIMIZATION OF COOPERATIVE MULTI-HOP WIRELESS SENSOR NETWORKS

6.1 Overview

In the previous sections, we developed the REACT protocol and shown its performances in various environments including energy-harvesting networks. Although REACT has shown its superior performance over non-CT protocols, it is still not clear whether the REACT protocol is an optimal way of using range-extension CT. One of the ways to determine whether a certain CT protocol can perform close to the optimal CT protocol or not is to obtain the optimal lifetime value when CT is used. This motivates us to investigate the optimal lifetime of cooperative routing, which is discussed in this chapter.

As we have mentioned in Section 2.3, the cooperative routing that relies on power-saving CT cannot guarantee the lifetime extension of multi-hop WSNs. Therefore, to obtain the true optimal lifetime or behavior of cooperative routing, both range extension and power saving of CT should be considered. Therefore, in this chapter, we study the optimal lifetime and behavior of cooperative routing that utilizes both features of CT. Also, as mentioned in Section 2.3, many CT-based routing works ignore the circuit energy consumption, which oversimplifies the problem or makes their approaches incorrect; in contrast, we try to correctly capture the energy cost of CT in our problem formulation.

By considering some unique characteristics of CT, we formulate a linear programming (LP) problem that can successfully capture the lifetime optimality of cooperative routing, which requires more sophisticated formulations than the case of non-CT. The LP solutions obtained from our formulation give the maximum achievable lifetime values of cooperative routing, and these values can be used as benchmarks to compare to the lifetimes of any cooperative routing protocols. We use this fact to compare the optimal lifetime of

CT with that of non-CT and the lifetime of the REACT protocol. Also, through evaluating/analyzing the LP solutions and performing network simulations for various cases, we identify important design parameters for the optimal cooperative routing protocol, which can help simplify the overall protocol design.

The remainder of this chapter is organized as follows. We formulate the lifetime-optimization problem of cooperative routing in Section 6.2. The evaluations of the formulated problem and network simulations for various cases are done in Section 6.3. In Section 6.3, we also analyze our LP solution to determine the key routing behaviors and design guidelines of cooperative routing. Finally, we summarize this chapter in Section 6.4.

6.2 Lifetime-Optimization Problem Formulation for Cooperative Routing

In addition to the common assumptions and definitions introduced in Section 2.4, we make the following assumptions and definitions. We consider a single-commodity[1] multi-hop wireless sensor network where sensed data is gathered at sink nodes. Each node can be a source except for the sink nodes, and the data generated by a source node is forwarded to one of the sink nodes. We also adopt some of implicit assumptions of [31] (the case of the lifetime-optimization problem for non-CT introduced in Section 2.1.2), such as collision is avoided so that no retransmission occurs, for our problem formulation. Nodes can adjust their transmit power, and a VMISO link can be formed between cooperating nodes and any node. An initiator uses *multicast* to share its data with its cooperators in the "CT sharing" or "multicast" phase followed by a VMISO communication.

We use the idea of [31] introduced in Section 2.1.2, and we use all the definitions made in Section 2.1.2. Note that, in our formulation, D is the set of sink nodes, and $D \subset E$. Through the entire section, instead of the terms "energy-constraint condition"

[1]Here, "single commodity" means that, when there are multiple sink nodes, a source node needs only to send its data to one of the sink nodes (a "single" destination).

and "data-conservation condition," we use the abbreviated terms Cond-EC and Cond-DC, respectively.

6.2.1 Intermediate Variables for CT

In this section, we define a set of intermediate variables, which capture the different ways of transmitting and receiving data in a CT-based network. These ways must be distinguished because they correspond to different values of energy consumption. In the next section, these intermediate variables will be expressed in terms of LP variables.

In a CT-based network, data can arrive at a node in different ways. It can be received over a SISO link from either a non-CT transmitter or from an initiator in the multicast (CT sharing) phase. Alternatively, it can arrive from a VMISO link or be self-generated. We denote the numbers of incoming data units arriving to Node i in each of these four ways by I_i^{n}, I_i^{m}, I_i^{v}, and I_i^{g}, respectively. Similarly, data can leave a node different ways. It can be transmitted over a SISO link as either a non-CT transmission or from an initiator in the multicast (CT sharing) phase of CT. Alternatively, the data can be transmitted over a VMISO transmission by a node, which can be either an initiator or a cooperator (non-initiator). We denote the numbers of outgoing data units from Node i in each of these three categories by O_i^{n}, O_i^{m}, and O_i^{v}, respectively. Here, O_i^{v} encompasses two cases where (i) Node i initiates and does CT and (ii) Node i has received a multicast (CT sharing) message from its neighboring initiator and does CT, which corresponds to I_i^{m}. Note that if Node i initiates x CT instances, then Node i will multicast x times (O_i^{m}=x) and x CTs, and therefore, the number of outgoing CT (VMISO) data from Node i initiated by Node i is equal to O_i^{m}. Therefore, the following holds:

$$O_i^{\mathrm{v}} = O_i^{\mathrm{m}} + I_i^{\mathrm{m}}. \tag{35}$$

6.2.2 Problem Formulation for CT

In this section, we define LP variables for CT and formulate the optimization problem using the variables. In our LP formulation, as in [50], we explain our problem in terms of data

packets, for convenience. However, the formulation can be explained in terms of bits or any other unit of data. For the remainder of this section, when necessary, we use the network in Figure 27 where solid lines indicate SISO links and dashed lines indicate VMISO links.

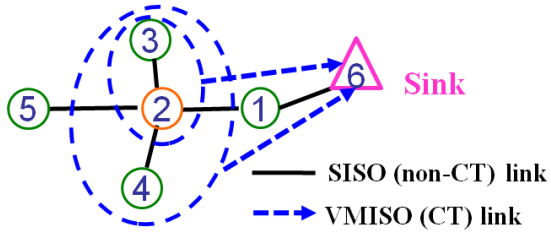


Figure 27: A network example. Node 2 can form a VMISO link between itself and Node 6 (sink node) by cooperating with (i) Node 3 or (ii) Nodes 3 and 4. Node 3 can also form a VMISO link to Node 6 by cooperating with Node 2. $N_c^{\max}$=3.

We first discuss why new LP variables are required for CT. Unlike the non-CT network, the energy consumption of a VMISO link between the initiator and a VMISO receiver highly depends on the combination of cooperating nodes. That is, as discussed in Appendix B, when transmit power control for CT is possible, the minimum required transmit power to successfully reach the VMISO receiver depends on the combination of cooperating nodes. Also, the energy consumption of the initiator for data sharing (multicast) depends on how it selects its cooperators (it only needs to reach all of its cooperators). Therefore, we need to define the new LP variables for CT that can represent the number of CT sharing messages from Node i to Node j for each possible combination of cooperating nodes.

We now define the new LP variables for CT. Let us denote by $R_{i,v}^N$, the set of N-tuples, $(r_1, r_2, ..., r_N)$, where $r_m \in S_i$ $(1 \leq m \leq N)$ and each tuple indicates the combination of cooperators that can reach Node v (a VMISO receiver) directly by cooperating with the initiator, Node i. For the example in Figure 27, $R_{2,6}^1$={(3)}, $R_{3,6}^1$={(2)}, and $R_{2,6}^2$={(3, 4)}. Note that, given v, there can be $N_c^{\max} - 1$ different $R_{i,v}^N$'s for Node i (in other words, $1 \leq N \leq N_c^{\max} - 1$). Now, let $n_{i,v,j}^{N,k}$ be the number of CT sharing data from Node i to Node j that is related to the k-th tuple in $R_{i,v}^N$, denoted by $R_{i,v}^{N,k}$. For example, when $R_{2,6}^2 = \{(3,4),(3,5)\}$, we can define variables $n_{2,6,3}^{2,1}$ and $n_{2,6,4}^{2,1}$ that are related to the first tuple (3,4), $R_{2,6}^{2,1}$, and variables

$n^{2,2}_{2,6,3}$ and $n^{2,2}_{2,6,5}$ are related to the second tuple (3,5), $R^{2,2}_{2,6}$. In this example, even though $n^{2,1}_{2,6,3}$ and $n^{2,2}_{2,6,3}$ both correspond to a SISO communication from Node 2 to Node 3, they are two distinct LP variables for CT. In addition to these variables, if we define $n^{\mathrm{nCT}}_{i,j}$ as the number of data units transmitted from Node i to Node j using non-CT ($n^{\mathrm{nCT}}_{i,j}$ matches with n_{ij} of the non-CT LP formulation in Section 2.1.2), we also have the variable $n^{\mathrm{nCT}}_{2,3}$ for the link between Node 2 and Node 3. Therefore, unlike the non-CT case, which requires a single variable n_{ij} for one SISO link from i to j, the LP formulation for CT has multiple variables defined for one SISO link. With $n^{N,k}_{i,v,j}$, we can correctly form Cond-EC for CT.

Given $R^{N}_{i,v}$ and fixed k, the following equality should hold for $j_m \in R^{N,k}_{i,v}$ $(1 \leq m \leq N)$ because CT sharing messages from Node i are multicast:

$$n^{N,k}_{i,v,j_1} = n^{N,k}_{i,v,j_2} = \ldots = n^{N,k}_{i,v,j_N}. \tag{36}$$

Now, we express the intermediate variables I's and O's in terms of the LP variables. For fixed v and N, the total number of outgoing data packets from Node i for the data sharing of CT is $(\sum_{j\in S_i}\sum_{k:j\in R^{N,k}_{i,v}} n^{N,k}_{i,v,j})/N$. The reason why the denominator N is required is because the data sharing is multicast so that $n^{N,k}_{i,v,j}$ for a fixed k cannot be considered as N different instances. Since $(\sum_{j\in S_i}\sum_{k:j\in R^{N,k}_{i,v}} n^{N,k}_{i,v,j})/N$ is for fixed v and N, we get O^{m}_i as follows:

$$O^{\mathrm{m}}_i = \sum_{N=1}^{N^{\max}_c-1}\left\{\left(\sum_{v:v\in A-i}\sum_{j\in S_i}\sum_{k:j\in R^{N,k}_{i,v}} n^{N,k}_{i,v,j}\right)/N\right\}. \tag{37}$$

Also,

$$I^{\mathrm{m}}_i = \sum_{N=1}^{N^{\max}_c-1}\left(\sum_{j:i\in S_j}\sum_{v:v\in A-j}\sum_{k:i\in R^{N,k}_{j,v}} n^{N,k}_{j,v,i}\right). \tag{38}$$

To formulate I^{v}_i, let us consider the case in Figure 27 and assume that Node 2 transmits one packet to Node 6 by cooperating with Nodes 3 and 4. Each of three nodes (Nodes 2-4) sends the same packet to Node 6, and total three data packets are sent for this VMISO communication. However, these packets are combined to be one packet at the sink node. Note that when Node 2 initiates x CT instances, Node 6 receives x packets regardless of how Node 2 selects its cooperators. This means that I^{v}_i of some Node i (e.g., Node 6 in

Figure 27) can be obtained from the number of outgoing CT packets "initiated" by some Node h (e.g., Node 2 in Figure 27). Since the number of outgoing CT packets "initiated" by Node h is equal to O_h^{m}, we can formulate I_i^{v} by (i) getting O_h^{m} for a fixed VMISO receiver i (using (37)) and (ii) using the fact that Node h can be any node except for the destination nodes and Node i, which results in

$$I_i^{\mathrm{v}} = \sum_{h:h\in A-D-i} \sum_{N=1}^{N_c^{\max}-1} \left\{ \left(\sum_{j\in S_h} \sum_{k:j\in R_{h,i}^{N,k}} n_{h,i,j}^{N,k} \right)/N \right\}. \tag{39}$$

For I_i^{n}, I_i^{g}, and O_i^{n}, we can simply use (4) and get

$$I_i^{\mathrm{n}} = \sum_{j:i\in S_j} n_{j,i}^{\mathrm{nCT}}, \quad I_i^{\mathrm{g}} = T \cdot Q_i, \quad O_i^{\mathrm{n}} = \sum_{j\in S_i} n_{i,j}^{\mathrm{nCT}}. \tag{40}$$

Let us consider Cond-DC for CT. In the case of CT, all incoming packets of Node i are transmitted to receiving nodes using either VMISO links or SISO links. Note that O_i^{m} is counting the packets for CT sharing, which is just an intermediate step to do CT and not the transmission to the intended receiver. Because of this, the sum of all incoming data packets should be equal to O_i^{n}+O_i^{v} (not O_i^{n}+O_i^{m}+O_i^{v}), which leads to the following Cond-DC of Node i for CT: $I_i^{\mathrm{n}} + I_i^{\mathrm{m}} + I_i^{\mathrm{v}} + I_i^{\mathrm{g}} = O_i^{\mathrm{n}} + O_i^{\mathrm{v}}$, or, using (35), $I_i^{\mathrm{n}} + I_i^{\mathrm{v}} + I_i^{\mathrm{g}} = O_i^{\mathrm{n}} + O_i^{\mathrm{m}}$.

Now, we formulate Cond-EC for CT. Note that the energy consumptions related to I_i^{n}, I_i^{m}, I_i^{v}, I_i^{g}, O_i^{n}, O_i^{m}, and O_i^{v} should all be considered for Cond-EC. The receiving energy consumption of SISO communication is e^{RX}, and it applies to I_i^{n} and I_i^{m}. The energy consumption of receiving CT packets depends on the number of cooperators, N_c (=N+1). For example, when the orthogonal diversity channel is obtained by using different time slots, a receiver has to spend more energy for VMISO reception than SISO reception because it has to receive all N_c packets. Therefore, if we define the energy consumption of receiving a CT packet (a CT data unit in general) when n cooperating nodes are transmitting as e_n^{RX}, the energy consumption related to I_i^{v} is

$$\sum_{h:h\in A-D-i} \sum_{N=1}^{N_c^{\max}-1} \left\{ \left(\sum_{j\in S_h} \sum_{k:j\in R_{h,i}^{N,k}} n_{h,i,j}^{N,k} \right) \cdot \frac{e_{N+1}^{\mathrm{RX}}}{N} \right\}. \tag{41}$$

For the transmit energy consumption, we define $e_{i,v}^{N,k}$ as the required transmit energy for Node i to send a data unit to Node v (VMISO receiver) when all elements (nodes) of $R_{i,v}^{N,k}$ and Node i cooperate. Also, we denote the required transmit energy for Node i to multicast a data unit to all nodes that are elements of $R_{i,v}^{N,k}$ by $E_{i,v}^{N,k}$. O_i^{v} is related to CT packets that are either initiated by Node i (related to O_i^{m}) or initiated by neighbors of Node i that Node i needs to cooperate (related to I_i^{m}). Note that when CT is initiated by one of Node i's neighbors, Node j, Node i has to use $e_{j,v}^{N,k}$ because the energy cost for doing CT depends on how Node j selects the rest of cooperators. Therefore, the energy consumption related to O_i^{v} is

$$\sum_{N=1}^{N_c^{\max}-1} \Bigg\{ \sum_{v:v\in A-i} \sum_{j\in S_i} \sum_{k:\, j\in R_{i,v}^{N,k}} \left(n_{i,v,j}^{N,k} \cdot e_{i,v}^{N,k} \right) / N + \sum_{j:i\in S_j} \sum_{v:v\in A-j} \sum_{k:i\in R_{j,v}^{N,k}} \left(n_{j,v,i}^{N,k} \cdot e_{j,v}^{N,k} \right) \Bigg\}. \tag{42}$$

e_{ij}^{TX} and $E_{i,v}^{N,k}$ are related to O_i^{n} and O_i^{m} respectively, and the energy consumption related to I_i^{g} can be expressed using the energy consumption for generating a data unit, denoted by e_i^{gen} (which is mostly the energy consumption for sensing in the case of the sensor network and is ignored in [31] and [50]). Cond-EC for Node i states that the total energy consumption of Node i should be less than or equal to its initial energy, and we can finally get the LP formulation for CT, which is summarized in Table 8.

When LP variables related to CT ($n_{i,v,j}^{N,k}$'s) are ignored, the LP formulation of CT in Table 8 becomes that of the non-CT. Also, (36) can reduce the number of required LP variables. Sink nodes require neither Cond-EC nor Cond-DC, and the number of packets that reach a sink node d can be calculated by I_d^{n}+I_d^{v}.

Note that when the orthogonal diversity channel is obtained by time [12], the cooperators can overhear the CT message of the initiator happening at the first time slot, and there is no explicit data-sharing phase for CT. This reduces the energy consumption of the initiator because it does not have to spend its energy for sharing the data. Therefore, in the cases where there is no explicit data-sharing phase, $E_{i,v}^{N,k}$=0 in Cond-EC, and Cond-DC

Table 8: Lifetime-optimization problem for cooperative routing.

Maximize T

s.t. $n_{i,j}^{\mathrm{nCT}} \geq 0$, $\forall i \in A, \forall j \in S_i$

$n_{i,v,j_1}^{N,k} = n_{i,v,j_2}^{N,k} = ... = n_{i,v,j_N}^{N,k} \geq 0$, $\forall i, v \in A, i \neq v, R_{i,v}^{N} \neq \{\}$

$$\sum_{j:i\in S_j} n_{ji}^{\mathrm{nCT}} + T \cdot Q_i + \sum_{h:h\in A-D-i} \sum_{N=1}^{N_c^{\max}-1} \left\{ \left(\sum_{j\in S_h} \sum_{k:j\in R_{h,i}^{N,k}} n_{h,i,j}^{N,k} \right)/N \right\}$$

$$= \sum_{j\in S_i} n_{ij}^{\mathrm{nCT}} + \sum_{N=1}^{N_c^{\max}-1} \left\{ \left(\sum_{v:v\in A-i} \sum_{j\in S_i} \sum_{k:j\in R_{i,v}^{N,k}} n_{i,v,j}^{N,k} \right)/N \right\}, \qquad \forall i \in A - D$$

$$\sum_{j:i\in S_j} n_{j,i}^{\mathrm{nCT}} e^{\mathrm{RX}} + \sum_{j\in S_i} n_{i,j}^{\mathrm{nCT}} e_{ij}^{\mathrm{TX}} + (TQ_i) e_i^{\mathrm{gen}} +$$

$$\sum_{N=1}^{N_c^{\max}-1} \left\{ \sum_{j:i\in S_j} \sum_{v:v\in A-j} \sum_{k:i\in R_{j,v}^{N,k}} n_{j,v,i}^{N,k} \cdot (e_{j,v}^{N,k} + e^{\mathrm{RX}}) + \right.$$

$$\sum_{v:v\in A-i} \sum_{j\in S_i} \sum_{k:j\in R_{i,v}^{N,k}} n_{i,v,j}^{N,k} \cdot (e_{i,v}^{N,k} + E_{i,v}^{N,k})/N +$$

$$\left. \sum_{h:h\in A-D-i} \left(\sum_{j\in S_h} \sum_{k:j\in R_{h,i}^{N,k}} n_{h,i,j}^{N,k} \right) \cdot e_{N+1}^{\mathrm{RX}}/N \right\} \quad \leq \quad E_i^{\mathrm{I}}, \qquad \forall i \in A - E$$

of CT remains the same because the intermediate CT sharing phase is already ignored in Cond-DC.

6.3 Evaluation and Analysis

In this section, we formulate and solve the lifetime-optimization problem of cooperative routing derived in Section 6.2 for given network models, and we analyze the LP results.

6.3.1 Simulation Models and Parameters

We consider multi-hop networks having a sink node with no energy constraint. All nodes have the same maximum transmission range of 20m ($d_{\text{tx}}^{\max}$=20m), and a SISO link exists from Node i to Node j if Node j is within $d_{\text{tx}}^{\max}$ from Node i. This makes SISO links deterministic. The LP solution[2] can provide the optimal lifetime bound of cooperative routing, and, to see the usefulness of this bound more clearly, we also perform network simulations for the REACT protocol introduced in Section 4.3 and compare its lifetime performance with the optimal case.

We use d_{req} in (52) for creating the set $R_{i,v}^{N}$ for Node i. That is, for Node i and Node v, we first fix N and consider each combination of Node i's cooperators and calculate d_{req}, and if (26) holds, we add the combination to $R_{i,v}^{N}$. We do this procedure for all possible N's ($1 \leq N \leq N_c^{\max} - 1$) and get the complete set of $R_{i,v}^{N}$. Based on $R_{i,v}^{N}$, the LP variables for CT are defined. We also use d_{req} to get $e_{i,v}^{N,k}$. That is, $e_{i,v}^{N,k}$ = 'circuit energy consumption' + 'radiated energy consumption required to reach d_{req}.' When calculating d_{req}, we use the diversity gains and path loss exponent defined in Section 2.4.2. We obtain the LP results by considering the *data packet* flows of the network, which makes the LP variables $n_{i,j}^{\text{nCT}}$ and $n_{i,v,j}^{N,k}$ the total number of "data packets" transmitted by Node i. The energy model is exactly the same as the one used in Section 4.4 (the energy model in [19]), and we assume 128 bytes of data. We do not consider the energy consumption for generating data ($e_{i,\text{gen}}$=0), and $E_{i,v}^{N,k}$=max($e_{ij_1}^{\text{TX}}$, $e_{ij_2}^{\text{TX}}$, ..., $e_{ij_N}^{\text{TX}}$) where j_1, j_2, ..., j_N are the elements of $R_{i,v}^{N,k}$. We measure the

[2]Note that, in LP, every local maximum is a global maximum.

lifetime performance in terms of the number of packets that successfully reach sink nodes. Since $n_{i,v,j}^{N,k}$ and $n_{i,j}^{\mathrm{nCT}}$ are integers, we formulate and solve a mixed-integer LP problem.

The orthogonal diversity channel is obtained by STBC [9], and the maximum number of orthogonal channels is three ($N_c^{\max}$=3). All nodes have the initial energy of 50mJ. Each node has identical data generation rate, and we set Q_i=1. Note that the conditions and assumptions made so far are for "our" evaluation purpose. One can use different link definitions, physical-layer characteristics, energy-consumption models, traffic models, etc. to evaluate one's own particular situation using the LP formulation derived in Section 6.2.

We consider square-shaped networks with a single sink node located at the bottom center of the network. There are 40 nodes in the network (excluding the sink node), and nodes are randomly deployed except for the sink node. One sample topology of 60m×60m networks is shown in Figure 28. For each size of network, 20 trials are performed (for both LP and the network simulations), and, in each trial, nodes are randomly relocated except for the sink node.

For simulating cooperative routing protocols such as REACT, we run network simulations to get the number of packets that reach the destination till one of the node dies. Note that, in the case of LP, we form matrices and vectors, and the LP formulation can be solved using any LP solution method, whereas, in the network simulation, packets are sent from node to node following routing decisions, and we measure the lifetime in terms of the number of packets that the destination receives. When simulating cooperative routing, as in Section 4.4, we consider two non-CT primary routing schemes: (i) AODV [56] and (ii) CMAX introduced in Section 2.1.3. Also, following Section 4.4, we denote REACT using AODV by REACT-AODV and REACT using CMAX by REACT-CMAX.

6.3.2 Optimal Network Lifetime of Cooperative Routing

In this section, we evaluate the optimal lifetime performances of cooperative routing, which can be obtained from the LP formulation in Section 6.2. These values are very important not only because they indicate how powerful cooperative routing can be when compared with

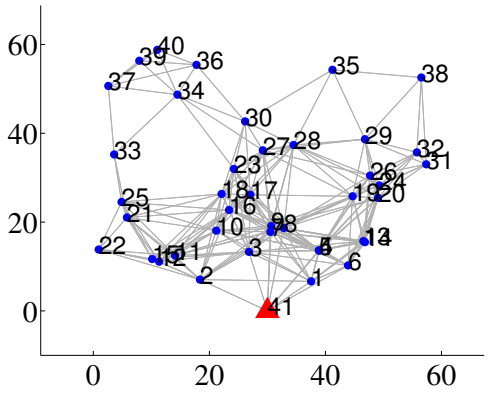


Figure 28: One sample topology of 60m×60m networks (solid lines indicate SISO links).

non-CT routing, but also because they can be used as a reference value when evaluating other cooperative routing methods. Note that, in our LP, maximizing T means maximizing the lifetime of the first node's death, and this translates into maximizing total number of generated packets, $T \cdot Q_i$, which maximizes the number of packets transmitted to the destination. Therefore, the performance bound of any cooperative routing can be obtained by solving LP and calculating the number of packets received by sink nodes.

REACT (and also PROTECT) considers the case of forming a VMISO link between cooperating nodes and the sink node (we will refer to this as "VMISO-Sink"), which means that there is only one VMISO receiver in the network. To see the benefits of allowing VMISO reception for any node, referred to as "VMISO-Any," we consider both VMISO-Sink and VMISO-Any when formulating LP. Note that in the LP formulation of VMISO-Any, one needs to consider e_n^{RX}, which does not need to be included for VMISO-Sink.

In Figure 29, the average lifetime performances (in terms of the number of packets) of four square-shaped networks are shown. In the figure, A, B, C, D, and E indicate optimal non-CT, optimal VMISO-Any, optimal VMISO-Sink, REACT-AODV, and REACT-CMAX, respectively. Note that LP is used for A, B, and C, and the network simulation is used for D and E. The circled solid lines indicate the mean values, and each dashed line is the outcome of one sample trial. When viewing the outcomes of individual trials, we

observe that the *relative* performance of the different schemes are highly correlated indicating the significant dependence of performance on topology, for any scheme. In other words, if Scheme A is better than Scheme B in one trial, A is very likely to be better than B in another trial. As can be seen from Figure 29, the optimal lifetime performances of CT clearly outperform those of non-CT. Figure 29 shows that VMISO-Any becomes notably better than VMISO-Sink when the network is large, and this is further discussed in Section 6.3.3. It can also be seen that REACT's performance is higher than that of the optimal non-CT scheme, however, it is observed that there is a notable gap between the optimal lifetime of VMISO-Sink and the lifetime of REACT (REACT is a VMISO-Sink scheme) especially for larger networks, showing that REACT cannot be considered as performing optimally. The lifetime optimization of cooperative routing, which uses VMISO links, gets more complicated as the network size grows because the increased number of hops from the sink node means that each sensed data may go through more than one CT decision (including whether to do CT or non-CT and selecting cooperators); all decisions need to be optimally made in order to maximize the network lifetime, and, because of the suboptimal methods of REACT, the lifetime of REACT deviates from the optimal as the network gets larger.

In summary, this section showed one important aspect of having the LP formulation for cooperative routing; it can determine whether a CT-based protocol behaves optimally or not, and one can determine the optimality of one's own cooperative routing protocol using the LP formulation developed in Section 6.2.

6.3.3 VMISO-Sink vs. VMISO-Any

This section discusses when and how VMISO-Any can be useful. It is clear that VMISO-Any can be advantageous over VMISO-Sink when it allows better energy usage compared to VMISO-Sink, and to explain those situations, we use simple 3-hop networks, one of which is illustrated in Figure 30a. For this network, we refer to the nodes i hops away from the sink node as "i-hop nodes," and the total number of i-hop nodes in the network is

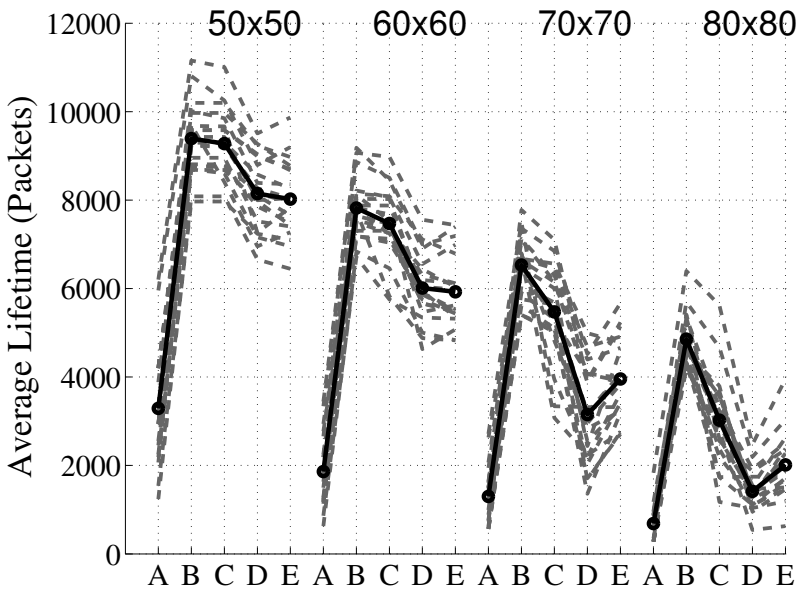
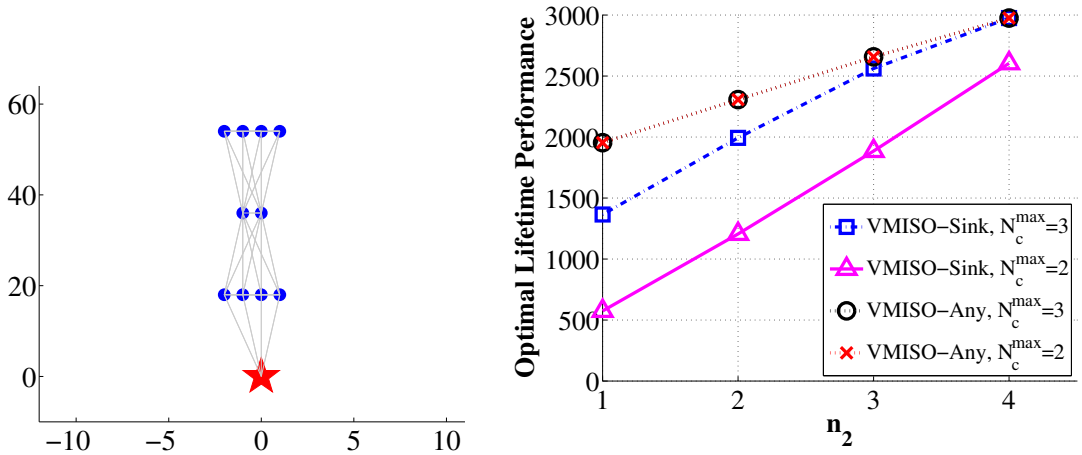


Figure 29: Average lifetime performance of four different network sizes (50m×50m, 60m×60m, 70m×70m, and 80m×80m) .

denoted by n_i. Here, $n_1 = n_3 = 4$, and n_2 varies. Also, for this network, we consider the case where a SISO link can be established between any i-hop node and any (i+1)-hop node.

When n_2 is very small (like one or two), then 2-hop nodes are the bottleneck nodes, and they die earlier than the others for non-CT networks, and, in order to prolong their lifetimes using CT, 3-hop nodes can use range-extension CT to reduce the relay burden of 2-hop nodes. Note that, based on the model in Section 6.3.1, VMISO-Sink requires at least three cooperating nodes to form a VMISO link from 3-hop nodes to the sink node, whereas, VMISO-Any has an additional option to form a VMISO link from 3-hop nodes to a 1-hop node using only two cooperating nodes, which can play an important role especially when 2-hop nodes are the bottlenecks. Moreover, if $N_c^{\max}$ is only allowed to be two, VMISO-Sink loses the option of forming a VMISO link using 3-hop nodes, which critically harms its lifetime performance. Also, since two cooperating nodes are enough for VMISO-any to form a VMISO link from 3-hop nodes to a 1-hop node, it is expected that increasing $N_c^{\max}$ will have no impact on the lifetime performance of VMISO-Any. The results in Figure 30b, which shows the optimal lifetimes of VMISO-Sink and VMISO-Any obtained



(a) A simple 3-hop network (solid lines indicate SISO links). n_1=4, n_2=2, and n_3=4.

(b) Lifetime performance. n_1 and n_3 are four, and n_2 (x-axis) is varied.

Figure 30: Comparisons of VMISO-Sink and VMISO-Any. Simple 3-hop networks.

from the LP formulation for $N_c^{\max}$=2 and 3, clarify the discussion made so far. We can see that two VMISO-Any cases are essentially identical as expected. When $N_c^{\max}$=3, VMISO-Sink approaches the performance of VMISO-Any as the bottleneck situation is resolved. However, when $N_c^{\max}$=2, it can be seen that the performance gap between VMISO-Sink and VMISO-Any is much worse than that of $N_c^{\max}$=3. This is because, for VMISO-Sink with $N_c^{\max}$=2, 3-hop nodes only do non-CT because three cooperating nodes are required to form a VMISO link to the sink node, and their energies cannot be fully utilized. Figure 30b shows that the existence of nodes in the network that cannot form VMISO link directly to the sink impacts the performance gap between VMISO-Any and VMISO-Sink more than the existence of the bottleneck. The discussion so far explains why the optimal performance of VMISO-Sink deviates from that of VMISO-Any in Figure 29 as the network area grows.

Figure 31 compares the average optimal lifetime performances of VMISO-Any and VMISO-Sink for the square-shaped topologies (nodes are randomly deployed) defined in Section 6.3.1 when $N_c^{\max}$=2 and 3. As can be seen from the figure, for small multi-hop networks requiring a few number of hops to reach the sink node, since VMISO-Sink can form VMISO link directly to the sink node with two cooperating nodes most of the time,

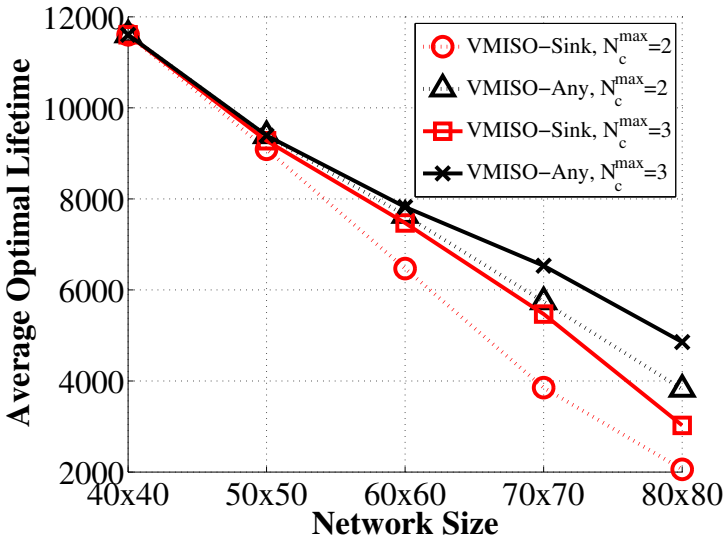


Figure 31: Average optimal lifetime performance of VMISO-Sink and VMISO-Any. Random square-shaped topologies (20 different network topologies per a fixed network size).

VMISO-Sink and VMISO-Any have almost identical performance regardless of $N_c^{\max}$. Note that using VMISO-Sink has the following advantages over VMISO-Any: (i) simpler protocol design because there is only one VMISO destination and (ii) simpler hardware for nodes other than the sink node because the diversity combining only needs to be done at the sink node, which makes VMISO-Sink more appropriate for simple WSNs than VMISO-Any. Also, having a smaller $N_c^{\max}$ simplifies the overall protocol because one needs to manage fewer orthogonal channels and fewer cooperators. Through the evaluation in this section, we have shown that reaching the near-optimal lifetime performance using VMISO-Sink with small $N_c^{\max}$ may be possible for small networks, which shows another usage of the LP formulation; one can use LP to determine (i) which VMISO scheme is desirable and (ii) $N_c^{\max}$ for given network and conditions, which can be helpful in simplifying a protocol.

6.3.4 Analysis of LP: Protocol Design Considerations for Cooperative Routing

The LP formulation derived in Section 6.2 can also provide observations regarding the optimal protocol behaviors of cooperative routing, which are discussed in this section. We analyze the LP results to learn important factors that may help us understand and design a

practical, near-optimal cooperative routing protocol. Here, we focus on the VMISO-Sink case because it is better suited for simple WSNs.

Let us consider all non-zero LP results related to Node 30 obtained from the network in Figure 28: $n^{\text{nCT}}_{30,18}$=9, $n^{\text{nCT}}_{30,27}$=57, $n^{\text{nCT}}_{36,30}$=150, $n^{1,1}_{30,41,17}$=221, and $n^{1,2}_{30,41,18}$=64. From these values, we can learn how Node 30 handles its packets when it decides to do non-CT or CT. That is, Node 30 transmits its packets to either 18 or 27 when it decides to do non-CT, and, when it decides to do CT, it uses only one cooperator (up to two cooperators are possible), and its cooperator is either Node 17 or 18. We denote the desirable cooperator groups for Node i that LP suggests by G_i (G_{30}= $\{\{17\},\{18\}\}$)[3]. Therefore, from LP, we can learn the non-CT route behavior and cooperator selection for maximizing the lifetime of cooperative routing. However, how a node decides whether to do CT or non-CT is not clear (in REACT, decision of doing CT is triggered by comparing the residual energies of nodes). Note that we know how many packets Node 30 decides to do CT, which is 285 (=221+64), and using the fact that total 351 packets are transmitted by Node 30, we can get the ratio of doing CT for Node 30, which is 0.812 (=285/351). Suppose we denote the ratio of doing CT for Node i by r^{CT}_i, and r^{CT}_i could somehow be known a priori. Then, in a practical protocol, Node i can decide to do CT (trigger CT) to match r^{CT}_i, and one of the simple ways to do this is to generate a random number between 0 and 1 and decide to do CT when the random number is less than or equal to r^{CT}_i. Likewise, when Node i decides to do CT, it has the ratio of selecting k-th group of G_i, which will be denoted by $r^{\text{coop}}_{i,k}$, as its cooperator (for example, $r^{\text{coop}}_{30,1}$=221/285). Similarly, when Node i decides to do non-CT, it has the ratio of selecting Node k as its next hop, which will be denoted by $r^{\text{nCT}}_{i,k}$. Again, one of the simple ways to match $r^{\text{coop}}_{i,k}$ is to generate a random number between 0 and 1 and select k-th group of G_i when the random number is between $\sum_{a=0}^{k-1} r^{\text{coop}}_{i,a}$ and $\sum_{a=0}^{k} r^{\text{coop}}_{i,a}$ where $r^{\text{coop}}_{i,0}$=0 (Node i can match $r^{\text{nCT}}_{i,k}$ using a similar method). $r^{\text{nCT}}_{i,k}$ is related to the non-CT route, G_i and $r^{\text{coop}}_{i,k}$ are related to the cooperator selection, and r^{CT}_i is related to CT decision

[3]The nodes far away from the sink node require more than one cooperator. For example, Node 38 in Figure 28 has G_{38}= $\{\{32, 35\}, \{31, 35\}, \{31, 32\}, \{29, 35\}\}$ according to the LP results.

(CT triggering). We are particularly interested in whether the ratio of doing CT (r_i^{CT}) can be effective when used as CT triggering because, unlike the non-CT route and cooperator selection that involve selecting neighbors, it is just a fixed ratio for Node i. Any scheme that tries to match $r_{i,k}^{\mathrm{nCT}}$, $r_{i,k}^{\mathrm{coop}}$, and r_i^{CT} will be referred to as *pseudo-optimal* because those values are obtained from LP (optimal), but the scheme only tries to match the ratio and is energy-unaware (not exactly optimal). The non-CT routing that tries to match $r_{i,k}^{\mathrm{nCT}}$ is referred to as pseudo-optimal routing (POR), and CT methods that try to match $r_{i,k}^{\mathrm{coop}}$ and r_i^{CT} are referred to as pseudo-optimal cooperator selection (POCS) and pseudo-optimal triggering (POT), respectively.

We perform network simulations to compare these pseudo-optimal schemes to identify the important factors in cooperative routing. For the network simulation, we use three non-CT routing algorithms for the primary routing of CT: (i) AODV, (ii) CMAX, and (iii) POR. The simulation parameters are exactly the same as the ones used in Section 6.3.2. We first see if POT can be useful in maximizing the lifetime. Figure 32 shows the average lifetime results (over 20 trials), obtained from network simulations, for the 60×60 networks used in Section 6.3.2. Figure 32 is showing four different combinations of CT triggering and cooperator selection: (i) original REACT, (ii) POT and POCS, (iii) POT and REACT's cooperator selection, and (iv) REACT's CT triggering and POCS. The dashed horizontal line in the figure indicates the average optimal lifetime obtained from the LP solution. As can be seen from Figure 32 (leftmost 3 bars), REACT with POR does not notably improve the performances of REACT-AODV and REACT-CMAX, which indicates that the primary routing scheme has little to do with the performance of REACT. When POT and POCS are used with POR or CMAX, it can be seen that the performance of the CT protocol improves greatly, and this means that the ratio of doing CT, when coupled with appropriate cooperator selections, can play an important role in extending the network lifetime. Also, as can be seen from Figure 32, using POT alone (or using POCS alone) worsens the performance of the original REACT, which indicates that the ratio of doing

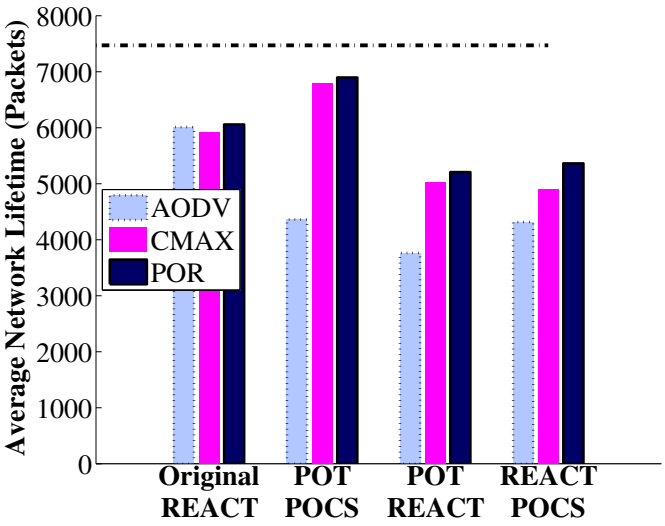


Figure 32: Average network lifetime of 60×60 networks.

CT and the cooperator selection should be jointly optimized.

Figure 33 shows the average lifetime results for the four different networks that we have considered in Section 6.3.2. Here, the average lifetime values are normalized by the optimal values obtained from LP. For the original REACT, Figure 33 shows the one with the best performance (out of the three possible primary routing cases). Here, we do not provide the cases where POT and POCS are not jointly used because they always perform worse than the original REACT protocol. Instead, we put the cases of using G_i only instead of $r_{i,k}^{\text{coop}}$ for the cooperator selection, referred to as POCS2 (that is, only cooperating group information is used without matching the rate $r_{i,k}^{\text{coop}}$). In POCS2, we need another method of choosing one of the multiple cooperating groups, and we consider the case where a node selects a cooperating group that has the largest minimum residual energy. From Figure 33, it can be observed that POCS2 with POR performs better than POCS with POR, which indicates that having desirable cooperator groups (G_i) is good enough and matching $r_{i,k}^{\text{coop}}$ is not as important as applying energy-awareness to G_i. When POT and POCS (or POCS2) are used with CMAX, denoted by POT,POCS(or POCS2)-CMAX, for "smaller" networks, the lifetime performance is not so different from that of POT and POCS using

POR (POT,POCS-POR), which is quite surprising considering the facts that (i) CMAX has nothing to do with LP results and (ii) $r_{i,k}^{\text{nCT}}$ of Node i influences the energy consumption of Node i's neighbors. Since nodes in smaller networks are only a few hops away from the destination, there are not many routing choices, and the energy-aware routing schemes such as CMAX try to evenly distribute the energy consumption of nodes as much as possible. The LP solution tries to fully utilize each node's energy, which eventually balances the energy consumption of nodes, and this can be the reason why both POR and CMAX, when supported by pseudo-optimal CT behavior, perform very well for smaller networks. This is very important because it implies that, for small networks, one may rely on an existing energy-aware routing scheme as the primary routing scheme in CT-based network, which eases the protocol design because jointly optimizing the primary routing scheme and CT behavior is not necessary. For larger networks, the performance of POT,POCS(or POCS2)-CMAX degrades compared to POT,POCS-POR, which means that CMAX does not perform optimally with POT and POCS when the network is large[4].

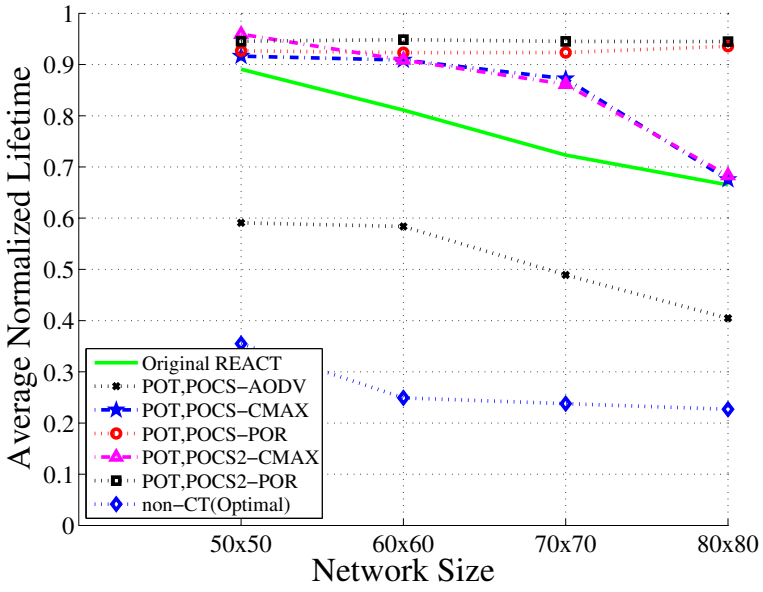


Figure 33: Average normalized lifetime results.

[4]The same conclusion can be drawn when different number of nodes are used for four different network sizes, and those results are omitted.

In summary, we identified two important factors, the ratio of doing CT (r_i^{CT}) and desirable cooperator selections (G_i), that one may consider when designing an optimal cooperative routing protocol, which may not need to be jointly optimized with the primary routing scheme for small networks. Note that, we have determined the values of G_i and r_i^{CT} using LP solutions, which must be computed "off-line" and require the data-generation rates to be known in advance. Whether one can design an online scheme[5] that can learn and use G_i and r_i^{CT} efficiently is an open question and is out of scope of this dissertation. If it is possible to determine G_i and r_i^{CT}, the protocol behavior of cooperative routing is straightforward; a node has to match r_i^{CT} for CT decision, and, when it decides to do CT, it uses one of cooperator groups in G_i to form VMISO.

6.4 Summary

In this chapter, we studied the optimal lifetime and behaviors of the cooperative routing. Using LP, we formulated the lifetime-optimization problem for cooperative routing that allows transmit power control and variable numbers of cooperators, which requires considerations of CT's unique characteristics and sophisticated variable definitions. By solving LP for cooperative routing, the optimal network lifetimes of cooperative routing were obtained and compared with those of the non-CT case to show the non-trivial lifetime improvement that CT can theoretically achieve. Also, we showed the usefulness of our LP by comparing the optimal performance to that of an online cooperative routing protocol through evaluations and simulations. We've also showed that allowing any node to be a VMISO receiver is not always necessary to achieve the optimal lifetime performance by providing the cases where allowing only the sink node to be a VMISO receiver is enough. We observed several key routing behaviors in the LP solutions, and, through testing protocols that approximate these behaviors, we determined certain factors that are important, which can be used when designing an optimal cooperative routing protocol. For small multi-hop networks requiring

[5]REACT and CMAX are online protocols that do not need to have the data generation rates in advance.

a few number of hops to reach the sink node, we found that the protocol design of cooperative routing can be simplified because one may use the sink node as a single VMISO receiver and rely on the existing energy-aware routing for the primary routing of CT.

CHAPTER 7

ONLINE COOPERATIVE ROUTING PROTOCOL FOR MAXIMIZING THE NETWORK LIFETIME

7.1 Overview

In the previous chapter, we have seen that there is a notable gap between the lifetime of REACT and that of the optimal case (VMISO-Sink). Motivated by this fact, in this chapter, we study the methods of designing an online cooperative routing scheme, and, through this study, we design an online cooperative routing scheme that is suitable for range-extension CT.

As we have mentioned in Section 2.1.1, online routing based on the minimum-cost route is one of the most popular online routing approaches and it is adopted in both CT [43], [49], [44] and non-CT [31], [35], [33], [34] works. One of the problems of the existing online routing methods that use the minimum-cost approach is that there is little or no theoretical justification for why the proposed methods should work well[1], and therefore, the proposed methods are justified only through running simulations and comparing with other existing methods. In contrast, in our approach, (i) several design criteria for online minimum-cost cooperative routing are defined, and, consequently, our online CT method can be justified theoretically using the criteria and (ii) some of the existing online minimum-cost routing approaches can be shown to not satisfy the criteria.

This chapter is organized as follows. Section 7.2 summarizes definitions, variables, and mathematical expressions required to design online cooperative routing. We define design criteria for online minimum-cost cooperative routing in Section 7.3, and based on these criteria we design online cooperative routing that is suitable for range-extension CT. We

[1]Note that, as we have seen in Section 2.1.3, [35] tries to theoretically justify the proposed method (CMAX) through competitive-ratio analysis, however, the analysis only provides the lower performance bound (in (6)) of the proposed method. If we use λ=100, which is the value used in 2.1.3 most of the time, $L(k) \geq 0.07 \cdot L_{opt}(k)$, which only guarantees that the performance of CMAX is larger than 7 percent of the optimal case.

also show that some of the existing online minimum-cost routing approaches are not so desirable in Section 7.3. Simulation results are provided in Section 7.5, and we summarize this chapter in Section 7.6.

7.2 Preliminaries

In this section, we give some assumptions and terminology, and then discuss how we calculate link and route costs.

Assumptions. The cooperative routing uses range-extension CT; a VMISO link is formed between cooperating nodes and a node that cannot be directly connected to using a SISO communication. The cooperative routing scheme is based on the minimum-cost approach; that is, a node assigns a link cost for each link according to the cost-metric function, and the route that gives the overall minimum cost is selected as the final route. Note that the minimum-cost approach is not only effective but also practical; the approach can be easily implemented in a distributed manner, as we have mentioned in Section 2.1.1. A node can adjust its transmit power, and all nodes have the same maximum transmission range, $d_{\text{tx}}^{\max}$. To simplify the analysis, we assume that a VMISO link can be formed only between cooperating nodes and a sink node (VMISO-Sink), which, as we have mentioned in Sections 3.2 and 6.3.3, allows WSNs to use simpler hardware for the nodes other than sink nodes.

Terminology. We now define terms, variables and mathematical expressions that are used throughout this chapter. We define the network lifetime as the number of data packets that successfully reach the sink node at the time of the first node's death. Let P^{C} be a set containing all possible cooperative routes between a given source and a given destination node. Also, we denote the i-th element of P^{C} by $P^{\text{C}}(i)$. For each path $P^{\text{C}}(i)$, we can define a set N_i that contains the node IDs of all the nodes (excluding the sink node) used in the path $P^{\text{C}}(i)$, and we can also define two subsets of N_i, which are (i) S_i, the set of node IDs of the nodes that are involved in SISO (non-CT) communications, and (ii) V_i, the set of node IDs

of the nodes that are involved in a VMISO (CT) communication ($S_i \bigcup V_i = N_i$). $P^{\mathrm{C}}(i)$ with $V_i = \emptyset$, $P^{\mathrm{C}}(i)$ with $V_i \neq \emptyset$, and $P^{\mathrm{C}}(i)$ with $S_i = \emptyset$ are referred to as the '*pure* non-CT path,' 'CT path,' and '*pure* CT path,' respectively.

Link and Route Costs. The lifetime of a node depends on its per-packet energy consumption, e, and its residual energy, r. When a node is a relay, e includes the receiving energy. We can define a cost-metric function as a function of r and e (at least), denoted by $g(r, e)$. Then, the total cost of using SISO links in $P^{\mathrm{C}}(i)$, denoted by C_i^S, can be expressed as

$$C_i^S = \sum_{k \in S_i} g(r_k, e_k(i)), \tag{43}$$

where r_k is the residual energy of Node k and $e_k(i)$ is the energy consumption of Node k when $P^{\mathrm{C}}(i)$ is used. Also, we denote the total route energy consumption for using $P^{\mathrm{C}}(i)$ by $e(i) = \sum_{k \in N_i} e_k(i)$. We denote the initial energy of Node k by E_k^{I}, and we use E^{I} to indicate the initial energy of an arbitrary node. The cost of using a VMISO link in $P^{\mathrm{C}}(i)$ will be denoted by C_i^V (discussed below), and the overall cost of using $P^{\mathrm{C}}(i)$, denoted by C_i, is

$$C_i = C_i^S + C_i^V, \tag{44}$$

and, we choose the route to be

$$P^{\mathrm{C}}(M) \qquad \text{s.t.} \quad M = \arg\min_i C_i. \tag{45}$$

Now, we consider the methods of calculating C_i^V. One way of getting C_i^V is to simply follow the form of (43), which gives

$$C_i^V = \sum_{k \in V_i} g(r_k, e_k(i)). \tag{46}$$

Note that (46) treats each cooperator's cost individually. Alternatively, one can treat multiple transmitters (cooperating nodes) in a VMISO communication as a single transmitter, which has an effect of converting a VMISO link cost to a SISO link cost. The similar approach has been considered in [43] and [44], where the authors considered the total energy

consumption of VMISO along with the minimum [43] or maximum [44] residual energy of cooperating nodes in their cost metric instead of considering each node's energy profile (although the authors did not explicitly state that cooperating nodes are treated as a whole). In this case

$$C_i^V = g(R(i), E^V(i)), \tag{47}$$

where $E^V(i)$=$\sum_{k\in V_i} e_k(i)$, which is the total energy consumed for the VMISO communication in $P^{\mathrm{C}}(i)$. The residual-energy part, $R(i)$, which depends on multiple residual energies of nodes, can also be treated as a whole, but there are multiple ways of getting $R(i)$, and we consider the following four cases that intuitively make sense: (i) the total residual energy, $R(i) = \sum_{k\in V_i} r_k$, (ii) the average, $R(i) = (\sum_{k\in V_i} r_k)/|V_i|$, (iii) the minimum residual energy, $R(i) = \min_{k\in V_i}(r_k)$, as in [43] and (iv) the maximum residual energy, $R(i) = \max_{k\in V_i}(r_k)$, as in [44]. We refer to these four cases as R^{sum}, R^{avg}, R^{min}, and R^{max}, respectively.

From the above discussion, there are two ways of obtaining C_i: C_i^S+(46) and C_i^S+(47), which will be denoted by C_i^{I} and $C_i^{\mathrm{W}}(R)$, respectively (I and W indicate that the cost of cooperating nodes is treated either "individually" or as a "whole," respectively), where R is one of R^{sum}, R^{avg}, R^{min}, and R^{max}. From now on, C_i^{I} and $C_i^{\mathrm{W}}(R)$ are referred to as the "cost-calculation method" of CT. For pure non-CT paths, $C_i^{\mathrm{W}}(R)$ and C_i^{I} are the same because C_i^V=0, and C_i=C_i^S.

We now introduce the concept of the obviously best route (OBR), which will be used in the following sections. Let us define the set N_i' as the set N_i (the set of Node IDs) sorted in an increasing order of the energy consumption of a node ($e_k(i)$). If there are multiple nodes with the same energy values in N_i, then, among those nodes, the Node IDs are sorted in an increasing order of the residual energy of a node. Also, we define the energy consumption and residual energy of k-th element of N_i' as $e^k(i)$ and $r^k(i)$, respectively. Then, we can define the OBR as follows.

Definition of OBR: Among all possible routes in P^{C}, an OBR is the path $P^{\mathrm{C}}(j)$ that satisfies $|N_j'| \le |N_i'|$, $e^k(j) \le e^k(i)$ and $r^k(j) \ge r^k(i)$ for all k, i, such that $1 \le k \le |N_j'|$ and

$1 \leq i \leq |P^C|$ $(i \neq j)$.

In other words, an OBR is the path that is the maximum-lifetime route choice of the moment because (i) it is the minimum-energy route and (ii) the minimum remaining energy of the nodes in the OBR after the nodes are used is always higher than or equal to that of the nodes in other routes (because $r^k(j) - e^k(j) \geq r^k(i) - e^k(i)$). Note that the OBR does not always exist, and also, there can be multiple OBRs[2].

7.3 Design Criteria for Cost-Metric Function and Cost-Calculation Method

In this section, we specify design criteria for the cost-metric function and cost-calculation method, which will be used for designing and determining desirable cost-metric functions and calculation methods. We first define three basic criteria that any online routing method should conform to in Section 7.3.1. Then, we introduce an additional criterion that applies to the case of range-extension CT in Section 7.3.2.

7.3.1 Basic Criteria

First, we give some general comments. The cost-metric function should have a form that selects the nodes that consume low energy and have relatively high residual energy. This can be achieved by making $g(r, e)$ a monotonically increasing function of e and a monotonically decreasing function of r; the cost of a link is lower than the other if the link in a route has (i) lower energy consumption or (ii) a node with higher residual energy, and this is our first criterion. For our next criterion, we consider the OBR. As we have mentioned in Section 7.2, the OBR does not always exist, however, if an OBR exists, the cost-metric function and cost-calculation method should guarantee that the OBR is selected. Finally, when all nodes in the network have the same residual energy, the cost-metric function and cost-calculation method should be able to select the path that gives the minimum total energy consumption so that the saved energy can be used for future traffics.

[2]Two paths $P^C(j)$ and $P^C(i)$ are the OBRs when $e^k(j) = e^k(i)$ and $r^k(j) = r^k(i)$ for all k's and $|N'_j| = |N'_i|$.

From the above discussions, we form the basic criteria for the cost-metric function and cost-calculation method as follows.

- **Criterion 1**: The cost-metric function, $g(r, e)$, should be a monotonically increasing function of e and a monotonically decreasing function of r.
- **Criterion 2**: The cost-metric function and cost-calculation method should always select an OBR if one exists.
- **Criterion 3**: When all nodes have the same residual energy, the cost-metric function and cost-calculation method should always select the path that minimizes the total energy consumption.

Note that Criterion 3 considers one particular case of the general case of the OBR in Criterion 2. That is, when all nodes have the same residual energy, the OBR is the minimum-energy route. However, Criterion 3 is simpler to check than Criterion 2, so it can be used when checking Criterion 2 is not straightforward. Note that, if the combination of cost-metric function and cost-calculation method does not satisfy Criterion 3, it does not satisfy Criterion 2.

7.3.2 Range-Extension-Specific Criterion

We now introduce Criterion 4, which applies to the case of range-extension CT only. As we have mentioned earlier, range-extension CT may avoid using the node that is about to die by establishing a new extended link when using the node cannot be avoided by non-CT or power-saving CT. In Criterion 4, given source and destination nodes, we denote the minimum-cost pure non-CT path among all possible pure non-CT paths by $P^{\mathrm{C}}(A)$ and the minimum-cost CT path among all possible CT paths by $P^{\mathrm{C}}(B)$.

- **Criterion 4**: If $P^{\mathrm{C}}(A)$ has a node that is to die if used, and if (i) $P^{\mathrm{C}}(B)$ can avoid using that node and (ii) the nodes involved in $P^{\mathrm{C}}(B)$ don't die after being used, then $P^{\mathrm{C}}(B)$ should always be used.

The case of Criterion 4 is illustrated in Figure 34 (showing only $P^{C}(A)$ and $P^{C}(B)$). In the figure, Node 1 in $P^{C}(A)$ will die if used, and using Node 1 can be avoided by using $P^{C}(B)$, which will use much more energy (more nodes), but the nodes in $P^{C}(B)$ won't die if used.

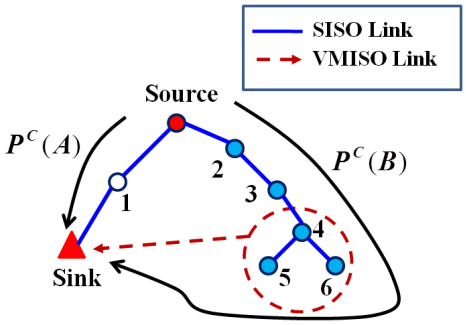


Figure 34: A network example illustrating Criterion 4. Node 1 in $P^{C}(A)$ will die if used.

7.4 Online Cooperative Routing: Design and Justification

In this section, we use the design criteria defined in Section 7.3 to determine desirable cost metrics and calculation methods of CT. We also design range-extension-specific cooperative routing using the criteria.

7.4.1 Desirable Cost-Metric Function

There can be many forms of $g(r, e)$ that satisfy Criterion 1, and because of this reason, we first determine the desirable form of $g(r, e)$ with justification. Since we want to extend the lifetime, let us consider the lifetime of a node. For a path $P^{C}(i)$, the lifetime of Node k in N_i, denoted by $T_k(i)$, is $r_k/e_k(i)$ (because the lifetime is measured by the number of data units that can be delivered). Since the minimum-cost approach finds the path that minimizes the overall cost, if we use $T_k(i)$ as its cost metric, the path will find the nodes with small lifetime. This problem can be solved if we use $T_k(i)^{-1}$ as the cost metric, which results in $g(r, e) = e/r$. This suggests that $g(r, e)$ that is proportional to e and inversely proportional r is a desirable choice, and this type of $g(r, e)$ satisfies Criterion 1. Since $g(r, e) = e/r$ gives no freedom of selecting the cost metric, we consider the form of $g(r, e)$=$e \cdot f(r)$ where $f(r)$

is a (strictly) monotonically decreasing function of r only ($f(r)$ should not depend on e or any node/link dependent variable), and, in the following section, we show that $e \cdot f(r)$ will satisfy Criterion 2 when an appropriate calculation method of CT is used.

Note that $g(r, e) = e/r$ (one case of $e \cdot f(r)$) happens to be one specific case of the FA algorithm introduced in Section 2.1.1 (x_1=1, x_2=1, and x_3=0). As we have mentioned in Section 2.1.1, x_2=x_3 is recommended by the authors of [31], which leads to the cost metric of the form $e_{ij}^{x_1}/(r_i/E_i^{\mathrm{I}})^{x_2}$. Note that the term r_i/E_i^{I} is the normalized residual energy (normalized by the initial energy of a node). Since the weighting factors (x_1 and x_2) have no reliable guidelines of selection and are not energy-related parameters[3], we simply ignore these, which leads to the form $e_{ij}/(r_i/E_i^{\mathrm{I}})$. Note that the form $e/(r/E^{\mathrm{I}})$ cannot be considered as one specific case of the form $e \cdot f(r)$ because $e/(r/E^{\mathrm{I}})$ not only depends on e and r but also E^{I}, which varies from node to node ($f(r)$ should be a function of the variable r only). In a later section (Section 7.4.3), we will justify using $g(r, e)$=$e \cdot f(r)$ instead of the form $e/(r/E^{\mathrm{I}})$.

7.4.2 Desirable Cost-Calculation Method for CT

Earlier, we defined five possible cost-calculation methods of CT: C_i^{I}, $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$. Here, we claim that C_i^{I} is more desirable than the others by proving the proposition below. We follow the proof with examples.

Proposition 1: The cost metric of the form $g(r, e) = e \cdot f(r)$ (defined in Section 7.4.1) with the calculation method of C_i^{I} can always satisfy Criterion 2, whereas the calculation methods $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ cannot.

Proof: Let us assume that, for given source and destination nodes, an OBR exists, and we denote this OBR by $P^{\mathrm{C}}(1)$. For Criterion 2 to be satisfied, any non-OBR, say $P^{\mathrm{C}}(a)$, should have higher cost, C_a, than the cost of using $P^{\mathrm{C}}(1)$, C_1. When the calculation method C_i^{I} is used, C_1=$\sum_{k=1}^{|N_1'|} e^k(1) \cdot f(r^k(1))$, and C_a=$\sum_{k=1}^{|N_a'|} e^k(a) \cdot f(r^k(a))$. When $|N_1'| < |N_a'|$, we

[3] [43] and [44] are some of the online routing works that use energy-related parameters only in the cost metric.

have $C_1 < C_a$ because, from the definition of the OBR, (i) $e^k(1) \cdot f(r^k(1)) \leq e^k(a) \cdot f(r^k(a))$ for all k's, where $1 \leq k \leq |N_1'|$ and (ii) C_a also has $e^k(a) \cdot f(r^k(a))$ terms for additional k's, where $|N_1'| + 1 \leq k \leq |N_a'|$. When $|N_1'| = |N_a'|$, from the definition of the OBR, $e^k(1) < e^k(a)$ or $r^k(1) > r^k(a)$ for at least one k because $P^{\mathrm{C}}(a)$ is not an OBR; we denote this k by $\bar{k}$. Then, C_1=$e^{\bar{k}}(1) \cdot f(r^{\bar{k}}(1)) + \sum_{k=1,k\neq\bar{k}}^{|N_1'|} e^k(1) \cdot f(r^k(1))$ and C_a=$e^{\bar{k}}(a) \cdot f(r^{\bar{k}}(a)) + \sum_{k=1,k\neq\bar{k}}^{|N_a'|} e^k(a) \cdot f(r^k(a))$, and, because (i) $e^{\bar{k}}(1) \cdot f(r^{\bar{k}}(1)) < e^{\bar{k}}(a) \cdot f(r^{\bar{k}}(a))$ and (ii) $\sum_{k=1,k\neq\bar{k}}^{|N_1'|} e^k(1) \cdot f(r^k(1)) \leq \sum_{k=1,k\neq\bar{k}}^{|N_a'|} e^k(a) \cdot f(r^k(a))$, we have $C_1 < C_a$. Therefore, $C_1 < C_a$ is always satisfied, and $P^{\mathrm{C}}(1)$, which is the OBR, is guaranteed to be selected.

Now, we show that $g(r, e) = e \cdot f(r)$ with $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ cannot satisfy Criterion 2. To do this, it is sufficient to show that $C_1 < C_a$ is "not always" guaranteed, and we do this by addressing a specific case where $C_1 < C_a$ is not satisfied. Let us consider the case where $|N_1'|$=$|N_a'|$, $e^k(1) = e^k(a)$ for all k's, $r^k(1) = r^k(a)$ for all k's except for k=$\tilde{k}$, and $r^{\tilde{k}}(1) > r^{\tilde{k}}(a)$; this case is referred to as 'Case A' in this section for convenience. When the OBR, $P^{\mathrm{C}}(1)$, is a pure CT path and $P^{\mathrm{C}}(a)$ is a pure non-CT path, $C_1^{\mathrm{W}}(R^{\mathrm{min}}) = g\left(R(1), \sum_{k=1}^{|N_1'|} e^k(1)\right) = f(R(1)) \cdot \sum_{k=1}^{|N_1'|} e^k(1)$, and C_a=$\sum_{k=1}^{|N_a'|} e^k(a) \cdot f(r^k(a))$. From the definition of N_i' and R^{min}, we have $R(1) < r^{\tilde{k}}(a)$ if $\tilde{k} \neq 1$ (if $\tilde{k} = 1$, $R(1)$=$r^{\tilde{k}}(a)$)[4], and, in this case, $C_1 > C_a$ (because $f(R(1)) > f(r^{\tilde{k}}(a))$, and the OBR is not selected.

In the opposite situation where the OBR, $P^{\mathrm{C}}(1)$, is a pure non-CT path and the arbitrary non-OBR, $P^{\mathrm{C}}(a)$, is a pure CT path, we can show that $g(r, e) = e \cdot f(r)$ with $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$ and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ cannot satisfy Criterion 2. In this situation, the OBR (pure non-CT) has C_1=$\sum_{k=1}^{|N_1'|} e^k(1) \cdot f(r^k(1))$, and the non-OBR (pure CT) has $C_a^{\mathrm{W}}(R) = g\left(R(a), \sum_{k=1}^{|N_a'|} e^k(a)\right)$ $= f(R(a)) \cdot \sum_{k=1}^{|N_a'|} e^k(a)$. If we consider Case A for R^{sum}, then $r^k(1) < R(a)$ for all k's, and therefore, $e^k(1) \cdot f(r^k(1)) > e^k(a) \cdot f(R(a))$ for all k's, and the OBR is not selected because $C_1 > C_a$. For R^{max}, if $\tilde{k} \neq |N_1'|$, then $r^{\tilde{k}}(1) < R(a)$, and the OBR is not selected because $C_1 > C_a$. Likewise, for R^{avg}, if $r^{\tilde{k}}(1) < \sum_{k=1}^{|N_a'|} r^k(a)/|N_a'|$ (which can happen from time to time), the OBR is not selected because $C_1 > C_a$.

[4]When $e^k(1) = e^k(a)$ for all k's, which is the case of Case A, from the definition of N_i', N_i' is sorted in an increasing order of the residual energy.

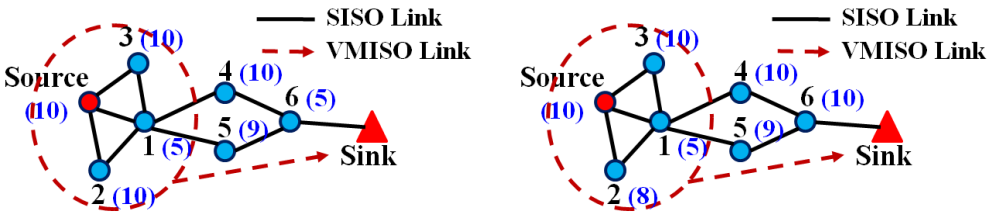
Therefore, $C_1 < C_a$ is not always guaranteed for $g(r,e) = e \cdot f(r)$ with $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$, so these calculation methods cannot satisfy Criterion 2. ■

Note that Case A in the proof is used to show a clear example, and, even without assuming Case A, it can be seen that there is no guarantee that $C_1 < C_a$ is satisfied for $C_i^{\mathrm{W}}(R)$. For instance, in the case of $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, we have $f(R(1)) \geq f(r^k(a))$ for "at least" one k, meaning that there is a possibility that $f(R(1))$ can be larger than $f(r^k(a))$, however, we need $f(R(1)) \leq f(r^k(a))$ for all k's to always guarantee $C_1 < C_a$.

Before moving on, we provide numerical examples showing that $g(r,e) = e \cdot f(r)$ with $C_i^{\mathrm{W}}(R)$ cannot always choose the OBR. We consider the network illustrated in Figure 35, where the node ID and residual energy (the number in parentheses) of each node are provided. For simplicity, we assume that each node in the route consumes a unit energy regardless of doing CT or non-CT[5]. For the example in Figure 35a, it is clear that the VMISO link formed directly to the sink node (in a dashed circle) is the OBR ($P^{\mathrm{C}}(1)$) and the second best route is the non-CT route, 'source→1→4→6→sink'. When $g(r,e) = e/r$ (which is one simple case of $g(r,e) = e \cdot f(r)$), the costs of the OBR and the second best route are $C_1^{\mathrm{W}}(R^{\mathrm{min}}) = 4/5$ (because $e(1) = 4$ and R^{min}=5) and 3/5, respectively, and therefore, the OBR is not selected (4/5 > 3/5). Note that when C_i^{I} is used, C_1^{I}=1/10+1/5+1/10+1/10=1/2, and the OBR is selected. Next, for the example in Figure 35b, the OBR is 'source→1→4→6→sink,' however, when $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ are used, the VMISO link formed directly to the sink node (which is not an OBR) is used because the OBR's cost, 1/2 (=1/10+1/5+1/10+1/10), is larger than the VMISO cost (4/33, 16/33, and 2/5, for $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$, respectively).

Note that since $C_i^{\mathrm{W}}(R)$ has the possibility of not choosing even the "obviously" desirable route, it is likely that it cannot find the close-to-optimal route from time to time. Therefore,

[5]Note that this assumption is to provide a simple example. Usually a non-CT source node requires one transmission, non-CT and CT relays require one reception and one transmission, and a CT initiator requires either one (when no explicit data sharing is required) or two transmissions (when explicit data sharing is required).



(a) Example 1. (b) Example 2.

Figure 35: Two network examples having the same topology (one of the possible cooperative routes is a pure CT path) with a different residual-energy distribution. The node ID of each node is provided on top or bottom of the node (except for the source node) along with the corresponding residual energy in parentheses.

it is desirable to use C_i^{I} for CT, and, in addition to the theoretical justification presented in this section, using C_i^{I} will be further justified through simulations in Section 7.5.3.

7.4.3 Normalized Residual Energy

As mentioned in Section 7.4.1, using the normalized residual energy in the link-cost metric is encouraged in [31], and this normalized residual energy appears or is adopted in some of the online energy-aware routing works as well [35], [49]. Here, we investigate using the normalized residual energy as in [31] in light of the criteria. When providing examples in this and the following sections, we use the simple network in Figure 36, where it is assumed that Node 3 is a source and there are only two possible paths to the destination: (i) forming a VMISO link to the sink node using Nodes 3 and 2, denoted by $P^{\mathrm{C}}(1)$, and (ii) a pure non-CT path via Node 1, denoted by $P^{\mathrm{C}}(2)$.

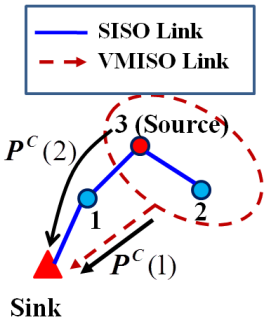


Figure 36: A network example. Only two routes exist from the source to the sink.

We now show that the cost metric of the form $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ is not desirable by proving the proposition below.

Proposition 2: The cost metric of the form $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$, even with the calculation method of C_i^{I}, cannot always satisfy Criterion 3, and thereby cannot always satisfy Criterion 2.

Proof: Since we consider Criterion 3, let us denote the equal residual energy by r_{eq}. Here we only consider the calculation method of C_i^{I} because the case where $C_i^{\mathrm{W}}(R)$ is used can be easily shown to be dissatisfying Criterion 3 by following the similar approach used for Proposition 1. Note that one particular case of $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$, where the initial energy of each node is the same for all nodes, falls into the case of $g(r, e)$=$e \cdot f(r)$, and, from Proposition 2, $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{max}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$ cannot satisfy Criterion 2. Let us consider the minimum-energy path, $P^{\mathrm{C}}(a)$, and an arbitrary non-minimum-energy path, $P^{\mathrm{C}}(b)$. C_a=$\sum_{k \in N_a} e_k(a) \cdot E_k^{\mathrm{I}}/r_{\mathrm{eq}}$ and C_b=$\sum_{k \in N_b} e_k(b) \cdot E_k^{\mathrm{I}}/r_{\mathrm{eq}}$, and to satisfy Criterion 3, we need $C_a < C_b$ so that $P^{\mathrm{C}}(a)$ is chosen. From the fact $P^{\mathrm{C}}(a)$ is the minimum-energy path, we have $\sum_{k \in N_a} e_k(a) < \sum_{k \in N_b} e_k(b)$, and, because of the term E_k^{I} in C_a and C_b, we may not have $C_a < C_b$. One simple example where we can observe $C_a > C_b$ is the case where $|N_a| = |N_b|$, E_k^{I}=$A/e_k(a)$ for all $k \in N_a$, E_k^{I}=$B/e_k(b)$ for all $k \in N_b$, and $A > B$ (A and B are positive constants); C_a=$|N_a| \cdot A$ and C_B=$|N_b| \cdot B$ so that $C_a > C_b$. Therefore, the form $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ cannot always satisfy Criterion 3, and, as we have addressed in Section 7.3.1, Criterion 2 is not satisfied when Criterion 3 is not satisfied. ■

In general, because of E_k^{I}, the node that has a "lower" initial energy is more frequently selected than the one with a relatively high initial energy, which is highly undesirable because the nodes that have less energy to spare in the beginning use more energy, which will make them die earlier than others. If we use the example in Figure 36 and assume that r_1=1, r_2=r_3=10, E_1=10, E_2=E_3=200, and $e_3(1)$=$e_2(1)$=$e_3(2)$=$e_1(2)$=1, even though $P^{\mathrm{C}}(1)$ should be chosen because Node 1 is to die if used, $P^{\mathrm{C}}(2)$ will be selected because C_1=20+20=40 > 30=20+10=C_2. Therefore, we do not recommend using the normalized residual energy.

If one wishes to use the normalized residual energy for some reason, one can use the $g(r, e) = e \cdot f(\bar{R})$ form instead, where $\bar{R}$=r/E_M^{I} and E_M^{I} is the maximum of the initial energies of all energy-constrained nodes in the network. Since (i) $f(\bar{R})$ does not depend on E_k^{I} and (ii) E_M^{I} is just a constant applied to all nodes, $f(\bar{R})$ is a function of residual energy only (that is, $f(\bar{R}) \subset f(r)$), and $g(r, e) = e \cdot f(\bar{R})$ will also satisfy Criterion 2 when used with C_i^{I}. The performances of the form $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ in [34] and $g(r, e) = e/(r/E_M^{\mathrm{I}})$ are compared in Section 7.5.2.

Note that $g(r, e) = e \cdot f(r)$ and $g(r, e) = e \cdot f(\bar{R})$ will choose exactly the same route when $f(r) = 1/r$ and C_i^{I} is used because if $e \cdot f(r)$ gives the relationship $C_a < C_b$ for two arbitrary paths, $e \cdot f(\bar{R})$ will give $E_M^{\mathrm{I}} \cdot C_a < E_M^{\mathrm{I}} \cdot C_b$ for the same paths.

7.4.4 Range-Extension-Specific Cost Metric

In this section, using Criterion 4, we propose a cost metric that is well suited for range-extension CT. We use the definitions of $P^{\mathrm{C}}(A)$ and $P^{\mathrm{C}}(B)$ made in Section 7.3.2 and Criterion 4. Also, we consider the calculation method of C_i^{I} in this section, and we assume that Node s is the source node and the node in $P^{\mathrm{C}}(A)$ that is about to die is Node a. For simplicity, we assume that (i) the circuit power is dominant so that $e_s(A) \approx e_s(B)$ and (ii) the power consumption of a node involved in non-CT relaying (consists of transmit and receive power) is identical to that of a node involved in a VMISO communication (because each cooperating node requires receiving and transmitting as in non-CT relaying); we denote this identical energy consumption by e_{eq}.

First, we consider the cost metric of the form e/r introduced in Section 7.4.1 and show that it cannot satisfy Criterion 4. We have C_A=$e_s(A)/r_s$+$\sum_{k \in N_A \setminus \{s\}} e_{\mathrm{eq}}/r_k$ and $C_B = e_s(B)/r_s + \sum_{k \in N_B \setminus \{s\}} e_{\mathrm{eq}}/r_k$, and we want $C_A > C_B$ so that $P^{\mathrm{C}}(B)$ can be selected. Now, to determine whether a cost metric can satisfy Criterion 4, we consider extreme conditions that make $C_A > C_B$ hard to satisfy. That is, since C_A is lowest when $|N_A \setminus \{s\}|$=1 (meaning only one relay exists in $P^{\mathrm{C}}(A)$ as in Figure 34), we set $|N_A \setminus \{s\}| = 1$, which leads to C_A=$e_s(A)/r_s$+e_{eq}/r_a, and we set r_a=e_{eq} indicating that Node a is to die if $P^{\mathrm{C}}(A)$ used. As a result, we get

C_A=$e_s(A)/r_s$+1. Also, since the second term of C_B is highest when r_k is small, we set r_k=2·e_{eq} for all $k \in N_B \setminus \{s\}$ indicating that Node k won't die if $P^{\text{C}}(B)$ used, which leads to $C_B = e_s(B)/r_s$+$|N_B \setminus \{s\}|/2$. Note that even if a cost metric cannot satisfy $C_A > C_B$ under the extreme conditions, it may satisfy $C_A > C_B$ when the conditions become less extreme. However, to truly claim that a certain cost metric can satisfy Criterion 4, the metric should satisfy Criterion 4 even in the extreme conditions. From the assumption that $e_s(A) \approx e_s(B)$, the first term in C_A and C_B cancels out, and we are comparing 1 (in C_A) with $|N_B \setminus \{s\}|/2$ (in C_B). When $|N_B \setminus \{s\}| > 2$ (this is likely because CT requires at least two cooperating nodes), we get $C_A < C_B$, which selects $P^{\text{C}}(A)$, and therefore, e/r with C_i^{I} cannot always satisfy Criterion 4.

Now, we design a cost metric that is better suited for range-extension CT. From the discussions of previous sections, $g(r, e) = e \cdot f(r)$ satisfies Criteria 2 and 3, and therefore, we want a cost metric of the form e·$f(r)$. Among many possible forms of e·$f(r)$, we propose $f(r)$=$(r/E_M^{\text{I}})^{-E_M^{\text{I}}/r}$, or equivalently $f(r)$=$\bar{R}^{-1/\bar{R}}$, which is carefully designed to satisfy both Criterion 1 and Criterion 4. Note that, unlike $f(r)$=$1/r$ where $1/r$ can be replaced by $1/\bar{R}$, $\bar{R}$ in $\bar{R}^{-1/\bar{R}}$ cannot be interchanged with r because it is required to satisfy Criterion 1. That is, the first derivative of $f(r)$=$(r/E_M^{\text{I}})^{-E_M^{\text{I}}/r}$ (with respect to r) is $E_M^{\text{I}} \cdot \bar{R}^{-1/\bar{R}} \cdot (\ln \bar{R} - 1)/r^2$, and since $\bar{R} \leq 1$, the derivative is always negative, which shows that $f(r)$ is a monotonically decreasing function of r (thereby satisfying Criterion 1). Also, using the assumptions and extreme conditions used for e/r, we can show that $g(r, e)$=$e \cdot (r/E_M^{\text{I}})^{-E_M^{\text{I}}/r}$ satisfies Criterion 4 as follows. Let us denote $e_{\text{eq}}/E_M^{\text{I}}$ by γ. If we ignore the cost of the source node (because it cancels out), then the effective costs are $\widetilde{C}_A$=$e_{\text{eq}} \cdot \gamma^{-1/\gamma}$ and $\widetilde{C}_B$=$e_{\text{eq}} \cdot (2\gamma)^{-1/2\gamma} \cdot |N_B \setminus \{s\}|$. It can be easily shown that $\widetilde{C}_A/\widetilde{C}_B = (2/\gamma)^{1/2\gamma}/|N_B \setminus \{s\}|$ so that comparing C_A and C_B is equivalent to comparing $\sqrt{2/\gamma}$ and $|N_B \setminus \{s\}|^{\gamma}$, respectively. Note that $\gamma \ll 1$ (because Node a is about to die), and $\sqrt{2/\gamma} > |N_B \setminus \{s\}|^{\gamma}$ is guaranteed to be satisfied because $\sqrt{2/\gamma}$ goes to infinity, whereas $|N_B \setminus \{s\}|^{\gamma}$ converges to one as γ reduces. For example, when γ=0.01 (1 percent of the maximum initial energy), $\sqrt{2/\gamma} \approx 14$ and $|N_B \setminus \{s\}|^{\gamma} \approx 1.1$ even with an

"unrealistically" large value of $|N_B \setminus \{s\}|$=10000 (note that $|N_B \setminus \{s\}|$ should be less than the total number of nodes in the network). Therefore, $C_A > C_B$ is satisfied (even under the extreme conditions), and Criterion 4 is satisfied for $e \cdot \bar{R}^{-1/\bar{R}}$.

Before we move on to the next section, we provide Table 9, which summarizes cost-metric functions and cost-calculation methods for the existing online routing works that are investigated in this chapter (for [31] introduced in Section 2.1.1, we set all the weighting factors to unity). Using the variables defined in this chapter, the cost metric in [43] can be expressed as $E^V(i)/\min_{k\in V_i}(r_k)$. Note that the cost metric in [43] includes the energy-related parameters of the VMISO receiver, however, in the VMISO-Sink case where the VMISO receiver is a sink node (which is considered in this chapter), this part can be ignored because the sink node is not energy constrained. Also, the cost metric in [44] can be expressed as $E^V(i)/\max_{k\in V_i}(r_k)$. The cost-metric functions in [43] and [44] have the form of $g(r, e) = e \cdot f(r)$, and therefore, according to Proposition 1, both cannot satisfy Criterion 2.

Table 9: Summary of the existing online routing approaches.

Reference	[31]	[43]	[44]
CT is considered.	No	Yes	Yes
Cost metric	$e/(r/E^{\mathrm{I}})$	e/r	e/r
Calculation method	C_i^S	$C_i^{\mathrm{W}}(R^{\min})$	$C_i^{\mathrm{W}}(R^{\max})$

7.5 Simulation

In this section, we justify our claims made in Sections 7.4.2 and 7.4.3, and we also see the performance of the range-extension-specific cost metric proposed in Section 7.4.4 through network simulations.

7.5.1 Simulation Models and Parameters

All nodes have the same maximum transmission range, $d_{\text{tx}}^{\max}$=20m. The energy model is the same as the one used in Section 4.4 (the energy model in [19]), and we assume 128 bytes of data. An initiator uses multicast to share its data with its cooperators in the "CT sharing" phase followed by a VMISO communication (therefore, an initiator requires two transmissions for one VMISO transmission). We use d_{req} in (25) and the condition in (26) to determine whether a group of nodes (cooperating nodes) can reach the sink node directly or not. When calculating d_{req}, we use the diversity gains and path loss exponent defined in Section 2.4.2. We assume that the orthogonal diversity channel is obtained by STBC and the maximum number of orthogonal channels is three.

The cost metrics and calculation methods discussed in this chapter are simulated, and we also provide the optimal solutions as well as the simulation results for the REACT protocol as references. As in Section 4.4, we consider REACT-AODV and REACT-CMAX. We measure the lifetime performance in terms of the number of packets that successfully reach sink nodes (till the first node dies). To compare the simulated results with the optimal performance of the off-line cooperative routing case, we use the lifetime-optimization problem for cooperative routing developed in Section 6.2 and obtain the theoretical lifetime performance. Note that the optimal performance is obtained by solving LP, not by running network simulations.

Source nodes are selected randomly, and each node has an equal chance of being selected as a source. We consider 70m×70m networks with a single sink node located at the bottom center of the network; an example is shown in Figure 37. There are 40 nodes in the network (excluding the sink node), and nodes are randomly deployed except for the sink node. 20 trials are performed, and, in each trial, nodes are randomly relocated except for the sink node. The node ID of a node is assigned according to its proximity to the sink node (the node closer to the sink node gets a lower node ID).

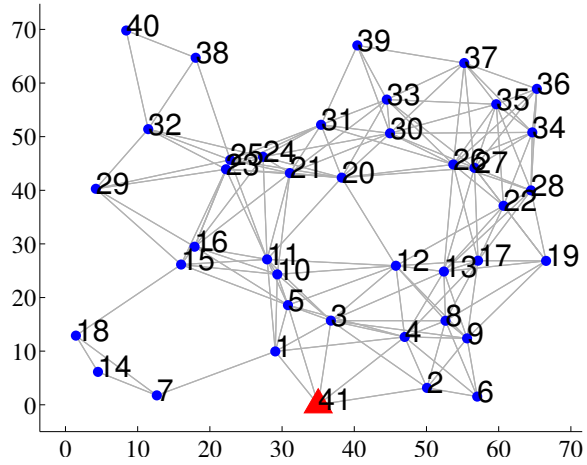


Figure 37: One sample topology of 70m×70m networks (solid lines indicate SISO links).

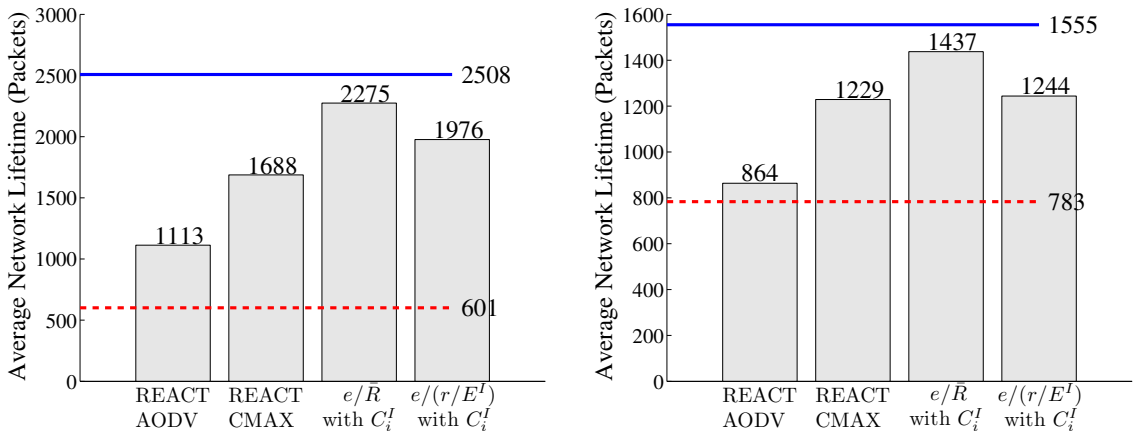
7.5.2 Simulation Results: Normalized Residual Energy

Here, we justify our claim in Section 7.4.3 that $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ is not recommended and $g(r, e) = e/(r/E^{\mathrm{I}}_M)$, or equivalently $g(r, e) = e/\bar{R}$, is better for the situation that nodes are initialized with unequal energies[6]. We consider two cases for the initial-energy distribution: (i) the nodes with even node IDs have 50mJ and the nodes with odd node IDs have 25mJ and (ii) Nodes 1-20 have the initial energy of 50mJ and Nodes 21-40 have 25mJ. Note that a node with a lower Node ID is closer to the sink node, and therefore, the second case of the initial-energy distribution considers the case where the nodes with high initial energies are strategically placed to mitigate the energy-hole problem.

Figure 38 shows the average lifetime performance for two cases considered; Figure 38a is Case (i) and Figure 38b is Case (ii). The straight line on top is the average lifetime of the optimal cooperative routing and the dashed line is the average lifetime of the optimal non-CT routing. As can be seen from both figures, $g(r, e) = e/\bar{R}$ outperforms $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$. Also, both $e/\bar{R}$ and $e/(r/E^{\mathrm{I}})$ perform better than REACT, however, in Case (ii), the performance of $e/(r/E^{\mathrm{I}})$ is close to that of REACT-CMAX, which is a very disappointing result for $e/(r/E^{\mathrm{I}})$ because (i) REACT relies on conventional non-CT routing and (ii) REACT does not need to figure out the minimum cost from a source to a sink node

[6]When all nodes have the same initial energy, $g(r, e, E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ and $g(r, e) = e/\bar{R}$ should perform exactly the same.

(in other words, REACT is much simpler than $e/(r/E^{\mathrm{I}})$). As explained in Section 7.4.3, $e/(r/E^{\mathrm{I}})$ tends to select the node that has a lower initial energy more frequently, and, in Case (ii), $e/(r/E^{\mathrm{I}})$ prefers using the nodes far away from the sink node (nodes with low initial energy) and the VMISO communication can be unnecessarily overused; REACT does not have this preference, and CT is not initiated by a node when the node's next-hop node in the primary route has more energy than itself. This is the reason why REACT, although it relies on a simple approach, can perform as well as $e/(r/E^{\mathrm{I}})$ in Case (ii). Also, since Case (ii) is trying to mitigate the energy-hole problem, we can observe that the average optimal lifetime of non-CT routing is relatively closer to that of cooperative routing than Case (i). In any case, $g(r,e) = e/\bar{R}$ outperforms both $g(r,e,E^{\mathrm{I}}) = e/(r/E^{\mathrm{I}})$ and REACT, which is consistent with our claim that using $e/(r/E^{\mathrm{I}})$ is less desirable than $g(r,e) = e/\bar{R}$. Note that, as mentioned earlier, the performance of $g(r,e) = e/\bar{R}$ is exactly the same as $g(r,e) = e/r$, and we also verified this using network simulations (results are omitted).



(a) Case (i): The nodes with even node IDs have the initial energy of 50mJ and the nodes with odd node IDs have 25mJ.

(b) Case (ii): Nodes 1-20 have the initial energy of 50mJ and Nodes 21-40 have 25mJ.

Figure 38: Average lifetime performances of $e/\bar{R}$ and $e/(r/E^{\mathrm{I}})$ when C_i^{I} is used.

Since the cost metric $g(r,e) = e/\bar{R}$ is insensitive to the differences in nodes' initial energies (by design), in the remaining sections, we consider only the case where all nodes have the same initial energy.

7.5.3 Simulation Results: C_i^{I} vs. $C_i^{\mathrm{W}}(R)$

Here, we justify our claim in Section 7.4.2 that C_i^{I} is more desirable than the four cases of $C_i^{\mathrm{W}}(R)$. We use the cost metric of the form $g(r, e) = e/\bar{R}$ (justified in Section 7.5.2), and we consider the case where all nodes have the initial energy of 50mJ.

Figure 39 shows the average lifetime performance of the five possible cost-calculation methods of CT that we consider: C_i^{I}, $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$. As can be seen from the figure, the calculation method C_i^{I} outperforms the rest. Among the four possible cases of $C_i^{\mathrm{W}}(R)$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$ performs best, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ performs worst. The worst performance of $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ can be intuitively understood by considering the fact that it only considers the maximum residual energy of all cooperating nodes; the nodes that are about to die can be selected as cooperating nodes as long as one of the cooperating nodes has a very large residual energy. Because of this reason, the performance of $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ is even worse than that of the worst case of the REACT protocol (REACT-AODV).

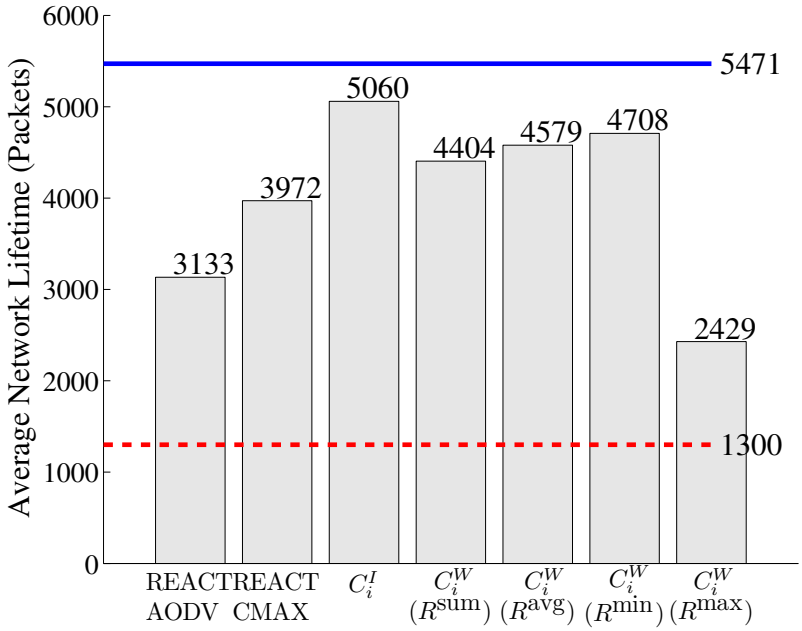


Figure 39: Average lifetime performances of C_i^{I}, $C_i^{\mathrm{W}}(R^{\mathrm{sum}})$, $C_i^{\mathrm{W}}(R^{\mathrm{avg}})$, $C_i^{\mathrm{W}}(R^{\mathrm{min}})$, and $C_i^{\mathrm{W}}(R^{\mathrm{max}})$ when $g(r, e) = e/\bar{R}$ is used.

From Figure 39, the average performance of C_i^{I} is more than 90 percent of the optimal, whereas $C_i^{\mathrm{W}}(R)$ has about 86 percent of the optimal at best ($C_i^{\mathrm{W}}(R^{\mathrm{min}})$), which backs up our

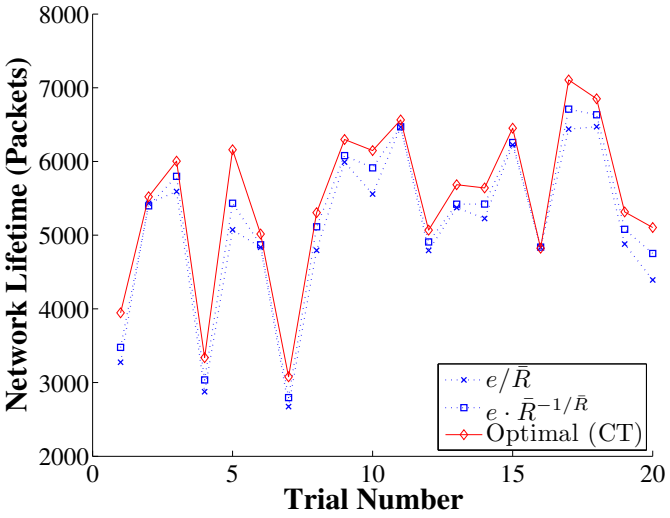


Figure 40: Lifetime performances of $e/\bar{R}$ and $e \cdot \bar{R}^{-1/\bar{R}}$ with C_i^{I}.

theoretical justification of using C_i^{I} discussed in Section 7.4.2.

7.5.4 Simulation Results: Range-Extension-Specific Cost Metrics

Here, we provide the network simulation results for the range-extension-specific cost metric presented in Section 7.4.4. Again, we consider the case where all nodes have the initial energy of 50mJ, and we use C_i^{I}.

Figure 40 shows the lifetime performance of the optimal (off-line) cooperative-routing case (obtained from LP, not network simulations), $e/\bar{R}$, and $e \cdot \bar{R}^{-1/\bar{R}}$ for each sample trial. As can be seen from the figure, $e \cdot \bar{R}^{-1/\bar{R}}$ outperforms $e/\bar{R}$ in most cases (except for the second trial), and, as a result, $e \cdot \bar{R}^{-1/\bar{R}}$ has higher average lifetime performance than $e/\bar{R}$; the average lifetime performances of the optimal case, $e/\bar{R}$, and $e \cdot \bar{R}^{-1/\bar{R}}$ are 5471, 5060, and 5220, respectively. The average lifetime of $e \cdot \bar{R}^{-1/\bar{R}}$ is 95 percent of the average optimal case, whereas $e/\bar{R}$ has around 92 percent of the optimal. As we have explained in Section 7.4.4, $e \cdot \bar{R}^{-1/\bar{R}}$ satisfies all four design criteria, whereas $e/\bar{R}$ cannot satisfy Criterion 4, and this can explain why the performance of $e \cdot \bar{R}^{-1/\bar{R}}$ is more close to the optimal than $e/\bar{R}$. Therefore, $e \cdot \bar{R}^{-1/\bar{R}}$ can be considered as more suitable for range-extension CT than $e/\bar{R}$.

7.6 Summary

In this chapter, we studied online minimum-cost cooperative routing that utilizes range-extension CT. By defining several criteria for designing minimum-cost routing, we theoretically justified our online approaches, which were verified through network simulations. Also, using the design criteria, we were able to find out that some of the existing minimum-cost approaches are not appropriate for cooperative routing. As for the cost-calculation method of CT, C_i^{I} turned out to be a desirable choice. In the case of the cost-metric function, the cost metric of the form $e \cdot f(r)$ defined in Section 7.4.1 satisfies all basic criteria, and two specific forms of $e \cdot f(r)$, $e/\bar{R}$ and $e \cdot \bar{R}^{-1/\bar{R}}$, were shown to perform very well most of the time when used with the calculation method of C_i^{I}. Therefore, the $e \cdot f(r)$ form with C_i^{I} can be a good candidate for online minimum-cost cooperative routing. Also, among all the cost metrics considered, the form $e \cdot \bar{R}^{-1/\bar{R}}$, which satisfies range-extension-specific criteria, was shown to have the best performance. However, as mentioned in 7.4.1, there can be many forms of cost-metric functions, and one can decide whether one's own approach is desirable or not by checking if the approach can meet the criteria introduced in this chapter.

CHAPTER 8

CONCLUDING REMARKS AND FUTURE RESEARCH

This research tackled one of the most important issues in multi-hop WSNs: extending the lifetime of the network. By considering the cooperative routing that mainly uses range-extension CT, this research showed that cooperative routing can mitigate the energy-hole problem and extend the network lifetime of multi-hop WSNs significantly.

Through this research, we provided useful tools that can approach the problems and solutions analytically and theoretically, which include the analytical model in Chapter 3, the formulation of lifetime-optimization problem in Chapter 6, and the design criteria in Chapter 7. Also, we designed several cooperative routing protocols that can extend the network lifetime in Chapters 3, 4, and 7. One important advantage of the cooperative routing protocols proposed in this dissertation is that they are very simple. That is, in order to decide whether to use a VMISO link or not, PROTECT requires only the level information obtained during the network initialization, and REACT requires only the information of the one-hop neighbors; both PROTECT and REACT can be built on top of an existing non-CT routing method and improve the network lifetime. Also, the minimum-cost cooperative routing we've explored in Chapter 7 can be easily implemented in a distributed manner using any existing shortest-path algorithm. Moreover, the proposed cooperative routing protocols consider the case of forming a VMISO link to a sink node only, which enables wireless sensor devices other than sink nodes to use simple hardware. Because the cooperative routing methods proposed in this research can efficiently utilize the energy of wireless sensor devices, they can also be used to provide better services than non-CT routing methods in EH-WSNs as we have shown in Chapter 5.

Because of the simplicity and nontrivial advantages of the proposed cooperative routing methods, they can be good candidates for future wireless networking techniques for WSNs. Also, the analytical tools developed in this research can be used by network operators to

analyze the performances of their multi-hop WSNs or may further be extended to solve more complicated problems. Some of the suggested future research topics are as follows:

1. Our lifetime-optimization problem for cooperative routing in Chapter 6 is formulated to maximize the lifetime of the first node's death. This objective may not be suitable for certain networks where some nodes can die because they do not critically harm network operations. Therefore, extending the lifetime-optimization problem to handle more flexible lifetime definition needs to be explored.

2. In Chapter 5, we have seen the advantages of cooperative routing over non-CT routing in EH-WSNs. Unlike the existing non-CT routing approach in [36], none of our proposed cooperative routing methods are harvesting aware. Therefore, modifying our existing methods or designing a new method of cooperative routing to better cope with the characteristics of EH-WSNs can be another research topic.

3. Although we have designed and fully justified the online cooperative routing methods in Chapter 7, we weren't able to determine whether the proposed methods are the best possible choices. Additional criteria or different approaches may be required to address the optimality of online cooperative routing, which is also a very interesting topic for the future research.

APPENDIX A

RANGE EXTENSION OF CT

Here, we explain how CT can extend the communication range of a wireless device, and the equations related to the range extension of CT are provided.

Consider the case where a node with a single antenna is transmitting. Let P_{TH} be the required received power that guarantees the SNR requirement to achieve a given level of performance at some desired data rate. The received power, denoted by P_{rx}, when the transmit power is P_{tx} and the distance between two communicating nodes is d_{link} is

$$P_{\mathrm{rx}} = k\frac{P_{\mathrm{tx}}}{d_{\mathrm{link}}^{\alpha}}, \tag{48}$$

where α is the path-loss exponent and k is a constant of proportionality [6]. (48) should be at least P_{TH} ($P_{\mathrm{rx}} \geq P_{\mathrm{TH}}$) to guarantee successful communication. Therefore, the minimum required transmit power, denoted by $P_{\mathrm{tx,min}}$, is $P_{\mathrm{TH}} \cdot d_{\mathrm{link}}^{\alpha}/k$.

Now, let us consider the case where CT [12] is used. Although each node has a single antenna, by using multiple physically-separated nodes (cooperating nodes), CT forms a "virtual array" and obtains diversity and array gains. Let us consider the case where the cooperating nodes are relatively close (compared with the destination far away from the cooperating nodes) to each other. In this case, each cooperating node has almost identical distance to the destination, and we denote this distance by d_{ct}. When N_c cooperating nodes use the same power $P_{\mathrm{tx,min}}$(=$P_{\mathrm{TH}} \cdot d_{\mathrm{link}}^{\alpha}/k$), the received power is

$$P_{\mathrm{rx}} = N_c \cdot k\frac{P_{\mathrm{tx,min}}}{d_{\mathrm{ct}}^{\alpha}} \cdot 10^{G/10} = N_c \cdot P_{\mathrm{TH}}\frac{d_{\mathrm{link}}^{\alpha}}{d_{\mathrm{ct}}^{\alpha}} \cdot 10^{G/10}, \tag{49}$$

where the factor N_c is the array gain and G is the diversity gain in dB. From (49), $N_c \cdot P_{\mathrm{TH}} \cdot 10^{G/10} \cdot (d_{\mathrm{link}}^{\alpha}/d_{\mathrm{ct}}^{\alpha}) \geq P_{\mathrm{TH}}$ should hold to guarantee successful communication, which leads to

$$d_{\mathrm{ct}} \leq d_{\mathrm{link}} \cdot (N_c \cdot 10^{G/10})^{1/\alpha} \triangleq d_{\mathrm{ext}}. \tag{50}$$

(50) shows us that CT allows a node to reach the distance up to d_{ext} when each cooperating node uses the power $P_{\text{TH}} \cdot d_{\text{link}}^{\alpha}/k$. Note that, with the power $P_{\text{TH}} \cdot d_{\text{link}}^{\alpha}/k$, a node that does not use CT can communicate with the nodes that are within d_{link} only, which is less than d_{ext}. Therefore, CT can extend the maximum transmission range of a node with a single antenna.

APPENDIX B

CALCULATION OF THE MINIMUM REQUIRED DISTANCE

Here, the minimum required transmit power that each cooperating node should use to reach a desired destination is explained. Since the transmit power of a transmitter is proportional to the distance that the transmitter can reach, the minimum required transmit power can be translated into the minimum required distance. In the derivation, the definitions made in Appendix A are used.

Note that we get (49) by assuming that the cooperating nodes are relatively close. To consider the case where the cooperating nodes are not relatively close and have different distances to the destination, we denote the distance between i-th cooperating node and the destination by d_i ($1 \leq i \leq N_c$). In this case, when N_c cooperating nodes use the same power $P_{\text{tx,min}}(=P_{\text{TH}} \cdot d_{\text{link}}^{\alpha}/k)$, the received power should satisfy

$$\sum_{i=1}^{N_c} P_{\text{TH}} \frac{d_{\text{link}}^{\alpha}}{d_i^{\alpha}} \cdot 10^{G/10} \geq P_{\text{TH}}, \tag{51}$$

which leads to

$$d_{\text{link}} \geq \left(10^{G/10} \cdot \sum_{i=1}^{N_c} d_i^{-\alpha} \right)^{-1/\alpha} \triangleq d_{\text{req}}. \tag{52}$$

(52) tells us that when each cooperating node uses the same transmit power $P_{\text{tx,min}} = P_{\text{TH}} \cdot d_{\text{req}}^{\alpha}/k$, CT can reach a fixed destination that is farther than d_{req}. In other words, it is enough for each cooperating node to use the minimum required transmit power $P_{\text{tx,min}} = P_{\text{TH}} \cdot d_{\text{req}}^{\alpha}/k$ to reach the desired destination. Without CT, a node can reach all nodes within d_{req} when the minimum required transmit power ($P_{\text{TH}} \cdot d_{\text{req}}^{\alpha}/k$) is used, and we refer to this d_{req} value as the "minimum required distance."

REFERENCES

[1] "Micaz datasheet." [Online]. Available: http://www.openautomation.net/uploadsproductos/micaz_datasheet.pdf

[2] *Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs).* IEEE Std 802.15.4™-2006.

[3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.

[4] V. Raghunathan, S. Ganeriwal, and M. Srivastava, "Emerging techniques for long lived wireless sensor networks," *IEEE Commun. Mag.*, vol. 44, no. 4, pp. 108–114, Apr. 2006.

[5] G. Anastasi, M. Conti, M. D. Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, May 2009.

[6] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed. Kluwer Academic Publishers, 2002.

[7] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[8] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, Inc., 2001.

[9] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[10] R. G. G. Dimitri P. Bertsekas, *Data Networks*. Prentice-Hall, 1987.

[11] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*. McGraw-Hill, 1996.

[12] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[13] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[14] G. J. Bradford and J. N. Laneman, "An experimental framework for the evaluation of cooperative diversity," in *Proc. 43rd Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2009.

[15] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[16] J. N. Laneman and G. W. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," in *Proc. IEEE GLOBECOM*, Nov. 2002, pp. 77–81.

[17] A. Stefanov and E. Erkip, "Cooperative space-time coding for wireless networks," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1804–1809, Nov. 2005.

[18] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.

[19] J. Li and P. Mohapatra, "An analytical model for the energy hole problem in many-to-one sensor networks," in *Proc. IEEE VTC*, 2005, pp. 2721–2725.

[20] S. Olariu and I. Stojmenović, "Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting," in *Proc. IEEE INFOCOM*, Apr. 2006.

[21] X. Wu, G. Chen, and S. K. Das, "Avoiding energy holes in wireless sensor networks with nonuniform node distribution," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 5, pp. 710–720, 2008.

[22] W. Wang, V. Srinivasan, and K.-C. Chua, "Using mobile relays to prolong the lifetime of wireless sensor networks," in *Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom)*, 2005.

[23] J. Lian, K. Naik, and G. B. Agnew, "Modelling and enhancing the data capacity of wireless sensor networks," *IEEE Monograph on Sensor Network Operations*, 2004.

[24] ——, "Data capacity improvement of wireless sensor networks using non-uniform sensor distribution," *Int'l J. Distributed Sensor Networks*, vol. 2, pp. 121–145, 2006.

[25] A.-F. Liu, X.-Y. Wu, and W.-H. Gui, "Research on energy hole problem for wireless sensor networks based on alternation between dormancy and work," in *Proc. The 9th International Conference for Young Computer Scientists*, 2008.

[26] M. Ettus, "System capacity, latency, and power consumption in multihop-routed ss-cdma wireless networks," in *Proc. IEEE Radio and Wireless Conf. (RAWCON'98)*, Aug. 1998, pp. 55–58.

[27] V. Rodoplu and T. H. Meng, "Minimum energy mobile wireless networks," *IEEE J. Select. Areas Commun.*, vol. 17, no. 8, pp. 1333–1344, Aug. 1999.

[28] T. H. Meng and V. Rodoplu, "Distributed network protocols for wireless communication," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS'98)*, Jun. 1998, pp. 600–603.

[29] S. Singh, M. Woo, and C. S. Raghavendra, "Power-aware routing in mobile ad hoc networks," in *Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom '98)*, Oct. 1998, pp. 181–190.

[30] J.-H. Chang and L. Tassiulas, "Routing for maximum system lifetime in wireless ad-hoc networks," in *Proc. the 37th Annu. Allerton Conf. Communication, Control, and Computing*, Sep. 1999.

[31] ——, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, Mar. 2000, pp. 22–31.

[32] Q. Li, J. Aslam, and D. Rus, "Online power-aware routing in wireless ad-hoc networks," in *Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom)*, Jul. 2001, pp. 97–107.

[33] C.-K. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Commun. Mag.*, pp. 138–147, Jun. 2001.

[34] J.-H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Trans. Networking*, vol. 12, no. 4, pp. 609–619, Aug. 2004.

[35] K. Kar, M. Kodialam, T. V. Lakshman, and L. Tassiulas, "Routing for network capacity maximization in energy-constrained ad-hoc networks," in *Proc. IEEE INFOCOM*, 2003.

[36] L. Lin, N. B. Shroff, and R. Srikant, "Asymptotically optimal energy-aware routing for multihop wireless networks with renewable energy sources," *IEEE/ACM Trans. Networking*, vol. 15, no. 5, pp. 1021–1034, 2007.

[37] F. Li, K. Wu, and A. Lippman, "Minimum energy cooperative path routing in all-wireless networks: Np-completeness and heuristic algorithms," *KICS/IEEE Journal of Communications and Networks*, vol. 10, no. 2, 2008.

[38] X. Fang, T. Hui, Z. Ping, and Y. Ning, "Cooperative routing strategies in ad hoc networks," in *Proc. IEEE VTC*, 2005, pp. 2509–2512.

[39] C. Pandana, W. P. Siriwongpairat, T. Himsoon, and K. J. R. Liu, "Distributed cooperative routing algorithms for maximizing network lifetime," in *Proc. IEEE WCNC*, Apr. 2006.

[40] A. E. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative routing in static wireless networks," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2185–2192, Nov. 2007.

[41] T. Himsoon, W. P. Siriwongpairat, Z. Han, and K. J. R. Liu, "Lifetime maximization via cooperative nodes and relay deployment in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 306–317, Feb. 2007.

[42] B. Maham, R. Narasimhan, and A. Hjϕrungnes, "Energy-efficient space-time coded cooperative routing in multihop wireless networks," in *Proc. IEEE GLOBECOM*, 2009.

[43] A. Aksu and O. Ercetin, "An energy-efficient routing protocol for networks with cooperative transmissions," in *Proc. IEEE ICC*, Jun. 2007, pp. 3345–3350.

[44] P. Zhou, W. Liu, W. Yuan, W. Cheng, and S. Wang, "An energy-efficient cooperative MISO-based routing protocol for wireless sensor networks," in *Proc. IEEE WCNC*, 2009.

[45] A. S. Ibrahim, Z. Han, and K. J. R. Liu, "Distributed energy-efficient cooperative routing in wireless networks," in *Proc. IEEE GLOBECOM*, 2007, pp. 4413–4418.

[46] ——, "Distributed energy-efficient cooperative routing in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3930–3941, Oct. 2008.

[47] G. Jakllari, S. V. Krishnamurthy, M. Faloutsos, P. V. Krishnamurthy, and O. Ercetin, "A cross-layer framework for exploiting virtual MISO links in mobile ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 579–594, 2007.

[48] S. Lakshmanan and R. Sivakumar, "Diversity routing for multi-hop wireless networks with cooperative transmissions," in *Proc. IEEE SECON*, 2009.

[49] Z. Han and H. V. Poor, "Lifetime improvement of wireless sensor networks by collaborative beamforming and cooperative transmission," in *Proc. IEEE ICC*, 2007.

[50] A. Sankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, 2004.

[51] C. R. Anderson and T. S. Rappaport, "In-building wideband partition loss measurements at 2.5 and 60 GHz," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 922–928, May 2004.

[52] Y. J. Chang, M. A. Ingram, and R. S. Frazier, "Cluster transmission time synchronization for cooperative transmission using software defined radio," in *Proc. IEEE ICC Workshop on Cooperative and Cognitive Mobile Networks*, May 2010.

[53] H. Jung, Y. J. Chang, and M. A. Ingram, "Experimental range extension of concurrent cooperative transmission in indoor environments at 2.4GHz," in *Proc. IEEE MILCOM*, 2010.

[54] M. U. Ilyas, M. Kim, and H. Radha, "Reducing packet losses in networks of commodity IEEE 802.15.4 sensor motes using cooperative communication and diversity combination," in *Proc. IEEE INFOCOM*, 2009.

[55] C. Zhou and B. Krishnamachari, "Localized topology generation mechanisms for self-configuring sensor networks," in *Proc. IEEE GLOBECOM*, 2003.

[56] C. E. Perkins, E. M. Belding-Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing," in *IETF RFC 3561*, Jul. 2003.

[57] C. Ó Mathuna, T. O'Donnell, R. Martinez-Catala, J. Rohan, and B. O'Flynn, "Energy scavenging for long-term deployable mote networks," *The International Journal of Pure and Applied Analytical Chemistry*, vol. 75, no. 3, pp. 613–623, 2008.

[58] W. S. Wang, T. O'Donnell, N. Wang, M. Hayes, B. O'Flynn, and C. Ó Mathuna, "Design considerations of sub-mW indoor light energy harvesting for wireless sensor systems," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 6, no. 2, 2010.

[59] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Aug. 1998, p. 156.

[60] "CC2420." Chipcon. [Online]. Available: http://focus.ti.com/lit/ds/symlink/cc2420.pdf

[61] S. K. Jayaweera, "Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 984–989, May 2006.

[62] Q. Wang, M. Hempstead, and W. Yang, "A realistic power consumption model for wireless sensor network devices," in *Proc. IEEE SECON*, 2006, pp. 286–295.

[63] X. Li, M. Chen, and W. Liu, "Application of STBC-encoded cooperative transmissions in wireless sensor networks," *IEEE Signal Processing Lett.*, vol. 12, no. 2, pp. 134–137, Feb. 2005.

[64] J. Li and P. Mohapatra, "Analytical modeling and mitigation techniques for the energy hole problem in sensor networks," *Pervasive and Mobile Computing*, vol. 3, no. 3, pp. 233–254, 2007.

[65] L. Bai, L. Zhao, and Z. Liao, "Energy balance in cooperative wireless sensor network," in *Proc. 14th European Wireless Conference*, 2008.

[66] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," in *Proc. ACM/IEEE Int. Conf. Mobile Computing and Networking (MobiCom)*, Aug. 2000.

[67] P. Levis, S. Madden, D. Gay, J. Polastre, R. Szewczyk, A. Woo, E. Brewer, and D. Culler, "The emergence of networking abstractions and techniques in TinyOS," in *1st Symposium on Networked Systems Design and Implementation (NSDI 2004)*, Mar. 2004.

[68] J. Luo and J.-P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proc. IEEE INFOCOM*, Mar. 2005.

[69] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. WSNA*, 2002, pp. 88–97.

[70] L. Simic, S. M. Berber, and K. W. Sowerby, "Energy-efficiency of cooperative diversity techniques in wireless sensor networks," in *Proc. IEEE PIMRC*, 2007, pp. 1–5.

[71] ——, "Partner choice and power allocation for energy efficient cooperation in wireless sensor networks," in *Proc. IEEE ICC*, 2008, pp. 4255–4260.

[72] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. IEEE INFOCOM*, 2002.

[73] J. Zhang, L. Fei, Q. Gao, and X.-H. Peng, "Energy-efficient multihop cooperative MISO transmission with optimal hop distance in wireless ad hoc networks," vol. 10, no. 10, pp. 3426–3435, 2011.

[74] Y.-W. Hong and A. Scaglione, "Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2844–2855, Oct. 2006.

[75] J. W. Jung and M. A. Ingram, "Lifetime optimization of multi-hop wireless sensor networks by regulating the frequency of use of cooperative transmission," in *Proc. IEEE MILCOM*, Nov. 2011.

[76] ——, "Proactive cooperative transmission to solve the energy hole problem in wireless sensor networks," *submitted to IEEE Transactions on Parallel and Distributed Systems*, Jul. 2011.

[77] ——, "Residual-energy-activated cooperative transmission (REACT) to avoid the energy hole," in *Proc. IEEE ICC Workshop on Cooperative and Cognitive Mobile Networks*, May 2010.

[78] ——, "Using range extension cooperative transmission in energy harvesting wireless sensor networks," *KICS/IEEE Journal of Communications and Networks*, vol. 14, no. 2, pp. 169–178, Apr. 2012.

[79] ——, "On using cooperative routing for lifetime optimization of multi-hop wireless sensor networks: Analysis and guidelines," *submitted to IEEE Transactions on Communications*, 2012.

[80] ——, "On the optimal lifetime of cooperative routing for multi-hop wireless sensor networks," *to appear in IEEE WCNC*, 2013.

[81] ——, "An RF channel emulator-based testbed for cooperative transmission using wireless sensor devices," in *Proc. NTMS 2008 Workshop on Wireless Sensor Networks: Theory and Practice*, Nov. 2008.

[82] J. W. Jung, A. Kailas, M. A. Ingram, and E. M. Popovici, "An evaluation of cooperation transmission considering practical energy models and passive reception," in *Proc. First International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, Aalborg, Denmark, Oct. 2008.

[83] J. W. Jung and M. A. Ingram, "Cooperative transmission range extension for duty cycle-limited wireless sensor networks," in *Proc. Wireless VITAE Conference*, 2011.

[84] M. A. Ingram, L. Thanayankizil, J. W. Jung, and A. Kailas, "Energy-harvesting wireless sensor networks," in *Globalization of Mobile and Wireless Communications: Today and in 2020*, Springer-Verlag, 2011.

VITA

Jin Woo Jung received the B.S. and M.S. degrees from Yonsei University (Seoul, South Korea) in 1999 and 2002, respectively. After graduating in 2002, he joined Datagate International (Seoul, South Korea), where he participated in security software R&D. In 2004, he joined Samsung Electronics (Suwon, South Korea), where he developed digital cable set-top boxes in the Telecommunication Network Division. In August 2007, he joined the Georgia Institute of Technology (Atlanta, Georgia, USA) as a PhD student. From August 2007 to December 2012, he worked under the tutelage of Prof. Mary Ann Ingram at the Smart Antenna Research Laboratory (SARL). With his research focusing on the energy-efficient communication protocols for wireless sensor networks, he received his PhD degree in Electrical and Computer Engineering from the Georgia Institute of Technology in Spring 2013. He worked for Apple, Inc. (Cupertino, California, USA) as a summer intern in 2010, and he is currently working at Panasonic Innovation Center (Atlanta, Georgia, USA).